The Dissertation Committee for Bridget Nicole Lee Certifies that this is the approved version of the following dissertation:

The Effect of Drama-based Instruction on PreK-16 Outcomes:

A Meta-Analysis of Research from 1985 – 2012

Committee:

Erika Patall, Supervisor

Marilla Svinicki

Stephanie Cawthon

Keisha Bentley-Edwards

Brent Hasty

The Effect of Drama-based Instruction on PreK-16 Outcomes:

A Meta-Analysis of Research from 1985 – 2012

by

Bridget Nicole Lee, B.S.Spe.; M.F.A

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2013

Dedication

The labor of this document is dedicated to my loving, supportive family. My husband, who is also pursuing his PhD, always finds ways to help me reconsider old ideas and make new connections. His love and support is endless. My two young children always make me laugh, even when I'm tired. My mom travels thousands of miles to cook, clean, and play with our kids. Above all, they give me the time to complain and celebrate the process and product of becoming a scholar of educational psychology and drama.

Acknowledgements

Writing “thank you” seems significantly short of the gratitude that I feel toward my mentors and advisers on this journey. I am happy and fortunate to consider them my academic colleagues and friends: Erika Patall, Stephanie Cawthon, Marilla Svinicki, and Kathryn Dawson. Erika took a chance on this crazy drama thing and supported me every step of the way. She allowed me to be in the process while at the same time, did not allow me to just sit idle. This is not an easy task, but she figured out how to do it. I could not have asked for a better mentor and friend in this process—from texting pictures of interview suits to endless discussions about our kids—I’ve loved it all. Stephanie shared many cups of coffee and let me “talk” through my ideas on whatever was handy—even on whiteboards and scratch paper. Her ability to give detailed notes and edits on a 30 page paper in a matter of minutes is unsurpassable. Even when she was overwhelmed, she acted excited to hear my knock on her door. Marilla, where do I even begin? Her unwavering loyalty and support for my work and for me was felt at every step in this process. Katie and I have shared many phases of our lives together: graduate school, becoming moms, developing programs, and becoming “academics”. I feel fortunate to have experienced all those journeys together and I know there are many more to come.

Additionally, I am indebted to the tireless work of my committee members, Brent Hasty and Keisha Bentley-Edwards, who have guided my process and understanding of this work. My fellow coders, Rebecca Steingut and Sean McNamara, for the endless coding work—it really did seem like it would never end! I cannot thank you enough for your work and dedication to this project. Every time you showed up to our meetings with one or two or five more documents coded, I was amazed and appreciative.

Finally, I need to thank my fellow graduate students who have listened to the complaints and celebrated the successes. What a road this is! Ariana Crowther, Jennifer Leach, Carlton Fong, Kadie Rackley, and Rebecca Steingut. Thank you for taking notes for me or grading papers for me when my kids were sick. Thank you for eating lunch at Chipotle when I was overwhelmed. Thank you for the conversations—not just the ones about *Q* between, but also the ones about our futures, being hangry, boy scouts, and Jennifer Lawrence vs. Anne Hatheway. Love to MERG!

The Effect of Drama-based Instruction on PreK-16 Outcomes:

A Meta-Analysis of Research from 1985 – 2012

Bridget Nicole Lee, PhD

The University of Texas at Austin, 2013

Supervisor: Erika Patall

In May 2011, the President's Committee on the Arts and Humanities (PCAH) released their report, "Reinvesting in Arts Education: Winning America's Future through Creative Schools." The PCAH heartily supported arts integration as an effective and cost efficient way to address teachers' and students' needs, referring to arts integration as "the most significant innovation in the field over the last two decades…" (2011). In the report, however, the PCAH called for a better understanding of the dimensions of quality and best practices—when, for whom, and what content areas are best served by arts integration methods. They called for research to clarify evidence on arts integration, not only as it relates to math and English standards but also as it relates to essential 21st century skills: creativity, critical thinking and collaboration (Partnership for 21st Century Skills, 2009).

One promising arts integration method is drama-based instruction. In the last twenty-five years, there have been numerous research studies assessing the effectiveness of integrating drama-based strategies into academic curriculum. However, we still do not

have a clear idea of the overall effectiveness of drama-based instructional strategies and, maybe more important, the conditions under which it is more or less effective. Prior research, including meta-analyses, have produced mixed results. Four previous meta-analyses found contradictory results and drew differing conclusions about moderators including gender and age. Much research has been published since the earlier meta-analyses. This, along with the need to clarify the effectiveness of DBI, warrants an updated review.

An initial search of the literature revealed 45 relevant studies since 1985, suggesting that there is indeed sufficient data to support an updated meta-analysis on the effectiveness of drama-based instruction for enhancing student academic achievement and other adaptive academic and social outcomes. The purpose of this meta-analysis is to provide a nuanced understanding of a specific arts integration method—drama-based instruction—and provide critical insight for practitioners regarding how it may be most profitably used in the classroom to benefit students.

Table of Contents

List of Tables

Chapter One Introduction

The current climate in primary and secondary schools is bleak. High school dropout rates escalate to nearly 50% in some high schools (National Center for Education Statistics, 2009). Budget crises force school districts to slash millions of dollars from already slim budgets. Teachers feel pressured to narrow curriculum to accommodate standardized testing (Policy, 2008). Given these conditions, the practical implications for educators are daunting. Teachers must find strategies to create relevant learning tasks that engage and motivate students of diverse backgrounds in the context of a distilled, standardized curriculum and with fewer resources. How might teachers meet these seemingly insurmountable challenges?

In May 2011, the President's Committee on the Arts and Humanities (PCAH) released their report, "Reinvesting in Arts Education: Winning America's Future through Creative Schools." The PCAH heartily supported arts integration as an effective and cost efficient way to address teachers' and students' needs, referring to arts integration as "the most significant innovation in the field over the last two decades…" (2011). In the report, however, the PCAH called for a better understanding of the dimensions of quality and best practices—when, for whom, and what content areas are best served by arts integration methods. They called for research to clarify evidence for arts integration, not only as it relates to math and English standards but also as it relates to essential 21st century skills: creativity, critical thinking, and collaboration (Partnership for 21st Century Skills, 2009).

One promising arts integration method is drama-based instructional strategies (DBI). The effectiveness of DBI was assessed in one comprehensive meta-analysis in 1986 and three subsequent limited meta-analyses in 1992 and 2000. However, a large literature on the effects of drama-based strategies on student's academic achievement and related academic outcomes has accumulated since then, making an updated synthesis particularly timely. Further, conflicting findings in existing primary research and previous meta-analyses suggest that we do not as of yet have a clear understanding of the overall effectiveness of DBI and, perhaps more importantly, the conditions under which it is more or less effective. In the present dissertation, I synthesize over twenty-five years of accumulated research on the effects of drama-based instruction on student outcomes to address the following overarching questions:

(1) What does the cumulative research suggest regarding the impact of drama-based instructional strategies on student outcomes including academic outcomes, psychosocial outcomes, and 21st century skills?

(2) Do characteristics of the intervention, students, or outcome influence the magnitude or direction of the effect of drama-based instructional strategies?

In the following chapter, I offer an overview of the conceptual and operational definitions of drama-based instruction, including its historical roots and current understanding in the field of drama and education. Then, I discuss the theoretical rationale for why drama-based instructional strategies may be expected to influence students' academic-related outcomes and what factors might be expected to explain the variability in the effects of drama-based instructional strategies. In particular, I focus on

five relevant theories of drama, learning, and motivation: pedagogy as performance, framework for differentiating instruction, experiential learning cycle, social constructivism, and self-determination theory, as well as highlight the empirical evidence that has, to this point, guided our understanding of drama-based instruction effects. Next, I discuss past seminal reviews and meta-analyses of drama-based instruction research and provide a rationale for conducting an updated research synthesis, including a meta-analysis, of the literature. Chapter three offers a detailed discussion of the methods for this study. In Chapter four, I present the quantitative results for the meta-analysis including: overall estimate of effects by outcome and moderator analyses of effects by outcome. Finally, Chapter five offers a discussion of these findings as well as their limitations. I conclude the report with implications for policy, practice, and future research based on the findings of this meta-analysis of the effects of drama-based instruction on Preschool through college across academic-related outcomes.

Chapter Two Literature Review

Tell me and I forget. Teach me and I remember. Involve me and I learn.

Drama-based instruction (DBI) is intended to provide a means through which a teacher may involve students in academic content and the classroom learning process. The purpose of this research synthesis is to examine the effectiveness of DBI on student outcomes through a synthesis and meta-analysis of the research. First, I present conceptual and operational definitions of DBI.

DEFINING DRAMA-BASED INSTRUCTIONAL STRATEGIES

What are drama-based instructional strategies? *Drama-based instructional strategies* (DBI) describe a collection of drama-based teaching and learning strategies lead by a facilitator to engage students in learning in both non-drama content (e.g., geometry, social/emotional skills, music, etc.) and drama content through a process rather than product oriented experience. In other words, DBI focuses on how students learn and engage with material rather than the final product or outcome of that learning. DBI strategies have roots in creative dramatics or in contemporary terms, creative drama, which has been defined as "an improvisational, non-exhibitional, process-centered form of drama in which participants are guided by a leader to imagine, enact and reflect upon human experiences" (Davis & Evans, 1982, p. 30). Although aspects of creative drama overlap with key ideas of DBI, creative drama does not specifically address the educative and/or psychosocial goals of DBI.

Specifically, the major defining features of DBI include that 1) it is facilitated and directed by a classroom teacher, teaching artist, or other facilitator trained in DBI, 2) it

works towards academic or psychosocial outcomes for the students involved, 3) it focuses on a process-oriented and reflective experience, and 4) it pools on the breadth and depth of applied theatre strategies.

The Role of the Leader. The leader may be a classroom teacher, a teaching artist, an arts specialist teacher or other facilitator trained in DBI. In DBI, the role of the leader lies on a continuum of involvement from participating in the DBI strategy to more of a guide on the side (King, 1993) or a facilitator for the learning. However, the leader always takes an active role in the drama work with the students as well as relates and reflects on the work as it connects to the non-drama content. He or she can guide the work of the students from the "sidelines" through providing clear instructions and side-coaching (Spolin, 1986). But beyond guidance, the leader intentionally co-constructs meaning of the material and experience with the students.

Goals of DBI. The goal of DBI is to influence student knowledge and skills both within the content domain of drama and sometimes more importantly, knowledge and skill development in other non-drama domains. The balance between the drama and non-drama knowledge and skill development goals can be thought of as a double helix where the drama and non-drama learning objectives are intertwined (Bowell & Heap, 2001) and teachers may choose to focus on one or the other outcomes at different times during the process. However, because DBI is intended for use in the context of non-drama coursework, teachers' may often preference the non-drama learning goals above drama skill development in order to serve their overarching curriculum objectives. In sum,

participants in DBI engage in learning non-drama and drama content through a DBI process.

Process-Oriented and Reflective Experience. Unique to drama-based instruction is a reliance on the tools of theatre and drama, but with the focus on learning as a process and reflective experience. For example, in a product-oriented experience, a leader may make choices based solely on aesthetics and audience understanding; whereas, in a process-oriented experience, a leader may make choices based on student learning and reflecting on that learning. Students may embody this learning through DBI, but there is no intention to have a formal presentation of that learning in a staged performance. For example, leaders may use a final culmination where students share what they have learned as a type of formative assessment but this sharing would not be a traditional play. In sum, DBI is a process-oriented and reflective experience for the leader and the students.

Applied Theatre Strategies. DBI strategies have their roots in applied theatre techniques. That is, DBI uses theatre-based activities and practices in an applied setting beyond a formal performance space. Theatre practitioners use various tools to create a production including, but not limited to: theatre games, improvisation, textual analysis, and character development. For example, in a theatre production, a Props Designer will read a script for clues about the type of objects (i.e., props) that need to be a part of the production. Each prop is carefully selected based on how it contributes to telling the relevant, appropriate story of the play. Translating this theater practice to DBI, a "prop" may be used in DBI to encourage the participants to infer information from the object or

artifact and think critically about the learning content. Another example is an actor's focus on contextual play analysis. An actor will read a line from the script and then based on his or her interpretation of the words, create acting objectives to play out in the scene. As this translates to DBI, a leader may provide a quote from a text at the beginning of a lesson. The students highlight and interpret the words that resonate with them. The leader encourages multiple interpretations and may even guide students to create a frozen image with their bodies to express the word's individual meaning. Just as the Prop Designer selects props and actors interpret lines of dialogue, the DBI leader uses the "prop" or "line of dialogue" to engage the participants in the curriculum and deepen their understanding and activate material. In this way, theatre processes and practices are applied to the educational setting for educative outcomes. See Table 1 for a summary of the components of DBI strategies. (For a detailed description of all the DBI strategies discussed in this document, see the unpublished document *Drama for Schools: A Handbook for using drama as an educational tool*.)

Component of DBI	Description
Role of the leader	• Actively participates in DBI with students • Co-constructs meaning of the material with the students • Relates experience to curriculum content
Goals of DBI	• Influence student knowledge and skills in drama and non-drama content
Process-oriented and reflective process	• Directs the experience in service to student learning and reflecting on that learning
Applied theatre strategies	• Uses theatre techniques and practices in the classroom • Engages students in the curriculum and deepens understanding of the material

Table 1. Description of DBI components

OVERLAPPING PRACTICES AND TERMINOLOGY

Broadly speaking, researchers and practitioners generally agree on what encompasses drama-based instruction. However, there are quite a few terms that have been used to describe activities that would fall within the category of drama-based

instruction. Researchers and practitioners have diversely referred to these techniques as creative drama (McCaslin, 1996), story dramatization (Ward, 1986), process drama (Heathcote & Bolton, 1995), drama-in-education (Bolton, Davis, & Lawrence, 1987), theatre-in-education (Jackson, 1993), theatre of the oppressed (Boal, 1974), applied theatre techniques (Cawthon & Dawson, 2009), theatre games (Spolin, 1986), enactment strategies (Willhelm, 2002), improvisation, and role playing. Many of these labels have developed through different contexts and have a specific intention. For example, theatre games were used to develop young actors, whereas enactment strategies began as a way to work with English Language Arts teachers. As a constellation of DBI strategies, they are all kinesthetic, interactive, and look like improvisation, storytelling, or role playing. In the following section, I offer a more detailed look at each of these terms as they relate to DBI.

Focused on a process-oriented drama experience, creative drama, story dramatization, drama-in-education, enactment strategies and process drama invite students to suspend their disbelief and enter into a magical world. Drama-in-education is the dominant term used in Great Britain and may be considered a broader term that encapsulates the other three terms. Creative drama, story dramatization and enactments originated in the U.S. and are typically more focused on a specific and structured story; whereas, process drama has an overarching structure but the drama may be very unstructured and directed by the students.

Theatre-in-Education (TIE) and Theatre of the Oppressed (TO) have a similar focus on the intersection of educational outcomes through a theatrical experience.

Typically, TIE starts with a preplanned, structured performance by professional teaching artists. Throughout the performance, however, the artists may stop the action and invite the observers to enact or respond to what is happening in the performance. Theatre of the Oppressed (TO) is rooted in the work of Augusto Boal, a Brazilian activist who fought to educate a largely illiterate population. In TO, the spect-actors (citizens) perform a theatrical event—many times unbeknownst to standers-by—in an effort to incite action of the observers. This activist type of theatre has been adapted and is used in classrooms in a more structured way.

Applied theatre in an educational setting is a broad term that encompasses DBI. As described earlier, these techniques may include all the theatrical elements of a traditional theatre production: script writing, directing, designing costumes, and improvising, and apply them to explore ideas or challenges outside of a theatrical space (e.g., a classroom, museum, or prison). Applied theatre is a broad term and in order to be considered DBI, the goals of applied theatre need to be educational in nature.

Some practitioners use a collection of specific terms for strategies to refer to the broad use of drama-based instruction. For example, improvisation is a specific type of DBI strategy but may be used interchangeably as a reference to the entire collection of DBI strategies. These specific terms include: improvisation, theatre games, role playing, and enactment strategies.

There is also an assortment of terms that describe drama activities that would not be considered drama-based instruction as I have defined it for this review, though they are likely to be confused with it. These include: dramatic play, theatre for young

audiences, and drama therapy. I offer a brief review of these terms in an effort to clarify the parameters of DBI in practice.

Dramatic play is typically defined as "the free play of very young children, in which they explore their universe, imitating the actions and character traits of those around them" (McCaslin, 1996). This type of play is sometimes called fantasy play, imaginative play, or creative play. All variations of this activity fall under the broad construct of play, that is, "activity that is not required but is enjoyed" (Play, 2001). Although process-oriented, the absence of a facilitator that intentionally guides learning prevents dramatic play from falling within the umbrella of DBI. If, however, a facilitator steps into the children's play and guides the creative play in a structured way, then this may be considered DBI.

Theatre for Young Audiences (TYA), Participatory Theatre, Children's Theatre, Storytelling, Youth theatre, Reader's theatre, ethno-drama, community-based theatre and other performative experiences involve an audience observing a performance by one or more actors/facilitators. TYA is defined as "The performance of a largely predetermined theatrical art work by living actors in the presence of an audience of young people" (Davis & Evans, 1982). Although many performances have an educative intention, the overarching purpose is to experience a quality, age-appropriate theatrical piece of work. These activities generally do not fall within the DBI umbrella because students rarely interact in the drama. If the observers are expected to enact or respond to the experience using a drama-based strategy, then these experiences may fall under the DBI umbrella.

Finally, the term Drama Therapy usually refers to enacting material "under the guise of symbolic dramatic play for therapeutic benefits" (Landy, 1983). This discipline borrows significantly from drama-based instruction and applied theatre techniques. However, the intention is for therapeutic benefits of clients. It may also be referred to as socio drama, playback theatre, psychodrama, or expressive therapy. See Table 2 for a summary of the terms that are used to describe DBI and non-DBI strategies.

DBI Strategies

Broad category of practice	Specific practices
Process-oriented drama-in-education	Creative drama
	Story dramatization
	Process drama
	Enactment strategies
Educational drama and performance	Theatre-in-education
	Theatre of the oppressed
Applied theatre	Applied theatre strategies
Specific strategies used to describe DBI	Improvisation
	Role playing
	Theatre games

Non-DBI Strategies

Broad categories of practice	Specific practices
Dramatic play	Fantasy play
	Imaginative play
	Creative play
Theatre for young audiences	Theatre for young audiences
	Participatory Theatre
	Children's Theatre
	Storytelling
	Youth theatre

Table 2. Cont. next page

	Reader's theatre
	Ethno-drama
	Community-based theatre
Drama therapy	Socio drama
	Playback theatre
	Psychodrama
	Expressive therapy

Table 2. Description of DBI and non-DBI terms

THEORETICAL UNDERPINNINGS OF DRAMA-BASED INSTRUCTIONAL STRATEGIES

In this section, I describe the foundational theories that I believe inform drama-based instructional strategies. The expected benefits of DBI are informed by the understanding of five theories: pedagogy as performance, framework for differentiating instruction, experiential learning cycle, social constructivism, and self-determination theory. At this point, a grand theory of DBI does not exist which is discussed further in the conclusions of this report. In this section, I highlight the relevant components of each theory, provide research evidence to support its claims, and offer specific examples of how DBI strategies are aligned with the theories.

Pedagogy as Performance

In part, performance studies theory interprets the idea of performance as educational or focused on growth and change (Schechner, 1993). That is, as we perform

our ideas, we, in turn, learn about our ideas. Over the last two decades of development of performance studies theory, educational researchers have examined the intersection of ideas between pedagogy and performance, creating the concept of pedagogy as performance (Pineau, 1994; Prendergast, 2008). This theoretical research suggests that educators who think of pedagogy as performance may consider and value educational play and process over procedures and products (Conquergood, 1989; Pineau, 1994).

By its very definition, DBI encourages the teacher to focus on improvisational and process-oriented strategies. By "performing" in the classroom, students may embody and/or kinesthetically respond to the curriculum in meaningful ways. In addition, teachers focused on the process of learning rather than the product thus encouraging deep rather than surface understanding of curriculum. Although this specific model of performance theory has not been explored through quantitative research, it seems that what we "do" as we perform our knowledge will likely impact what we "know".

Based on this theory, DBI would be expected to enhance learning outcomes compared to traditional approaches because it allows for a performance of knowledge to deepen learning. To that end, I offer two connected examples of how students "perform" their ideas through specific strategies: *Role on the Wall* and *Paired Improvisation*. As students are learning about a main character in a book, the teacher draws an outline of the character on the board. Then ask students to write all the other characters and messages that may be influencing the main character's decisions. Next, the teacher asks the students to write on the inside of the drawing how these characters and their messages are making the main character feel. Finally, through *Paired Improvisation,* the teacher invites

the students to simultaneously perform a dialogue between the main character and one of the other characters. In this way, the students embody their ideas from *Role on the Wall* and extend them for a deeper understanding.

Framework for Differentiating Instruction

In thinking about why we might expect DBI to effectively enhance learning and learning-related outcomes of the student, it is helpful to consider where DBI strategies lie in the context of the conceptual framework for differentiating instruction offered by Chi (2009). Based on her review of research, Chi (2009) suggests instruction can be categorized into four major types: passive, active, constructive, and interactive. Chi (2009) conceptualizes the various types of instruction as distinguishable both in terms of the structure of the activities, as well as in terms of the observable student behaviors and cognitive processes that occur during each type of activity.

During passive instruction, learners are not physically active nor are they attending to the relevant material. In active instruction, learners are actively participating in some way (e.g., fixating their gaze on relevant material, searching prior knowledge for connections). In constructive instructional strategies, learners make inferences that go beyond the given information, for example, by explaining their reasoning or making new connections between ideas. According to Chi, the final instructional type is interactive. During interactive strategies, students dialogue extensively about the topic and incorporate one another's ideas, building on a partner's ideas, revising personal misconceptions, or challenging and/or defending ideas, among other activities.

Past research suggests that active strategies are more effective for student learning than passive strategies. For example, in a lab experiment, one group of students were given a text to study for a quiz and another group of students were given the same text to study for a quiz but also prompted to explain what they were studying. The group that was required to explain their studying scored significantly better on the quiz (Williams & Lombrozo, 2010). This seemingly simple strategy of self-explaining shifted the learning strategy from passive to active and suggests that the learner was able to attend more effectively to the material.

Additional research shows that constructive strategies are more effective for learning outcomes than active strategies. In a study that tested mapping skills with fourth graders, a group of students completed a mapping activity in which the baseline group had to match flags to a corresponding model of a map. In the other condition, the small group of students had to match the flags to the corresponding model, but in addition, they also had to write down the reasons and the clues they used for making the decision to place flags on the model (Kastens & Liben, 2007). In this study, the students using the constructive strategies where they had to make explicit their connections and reasons for placing the flags scored significantly better on a subsequent test of map-making skills. This suggests that students learn more effectively when they are not only active, i.e., placing the flags, but also when students discuss their reasoning to justify their activity.

Finally, research shows that interactive strategies are more effective for learning than constructive strategies. For example, in a study by Roscoe and Chi (2007a), one condition of students were given the task to create a tape-recorded lesson to teach a

fellow student about the anatomy and inner workings of the eye. They were prompted to clarify any confusing information through specific examples in the lesson they created. This activity can be classified as constructive because the student needs to generate information beyond what was given and make new connections to facilitate the creation of a lesson. In the other condition, students were given the task to teach the content to a "learning" student who was in the room with them. This student was instructed to ask questions when they did not understand the material, therefore shifting the activity to an interactive strategy. The substantive dialogue between the teacher/student and the learning student built a new understanding and challenged the teaching students' conceptions of the topic. The students who taught the material to the learning student performed significantly better than the students who created a tape-recorded lesson to teach the material (Roscoe & Chi, 2007a). This suggests that students learn more effectively when they construct, reconsider, and construct again their understanding of a concept in relation with another student.

In order to map DBI onto this framework, I offer an example of three DBI strategies that can be classified as active, constructive, or interactive. Due to the active nature of DBI, it is assumed that none of the DBI strategies would be classified as passive. As an example of an active strategy, the theatre game *The Truth about Me* could be played as a review for the mathematical properties of integers. Each student is assigned a number (e.g., 1-20). One student says, "The truth about me is that I am a prime number." Then all the students who have prime numbers exchange places in the circle. One student will remain in the middle and continue with a truth statement about his or her

integer. This review is physically active—learners need to attend to relevant information to know when to move; however, learners will likely be using their prior knowledge rather than creating new connections.

To illustrate a constructive strategy, in *Artifacts,* the teacher presents an object or artifact that the students have not seen previously. The teacher invites students to describe the details of what is observable about the object. Then the teacher asks the students to infer what they think those details might mean. Finally the teacher invites the students to make predictions about the objects' origin and owner. Typically, this activity is not very physically active as students discuss the object. However, the students engage in deep cognitive processing as they make new connections between what is seen and what it could mean. In particular, this activity allows students to justify each prediction with observable characteristics.

Finally, for an example of interactive strategies, *Real/Ideal Images* is a DBI strategy that invites the students to create two frozen images of a concept in a "real life" version and an "ideal" version. The teacher uses the similarities and differences in the images to broaden and deepen students understanding of a concept. The discussion between the students and the teacher builds upon and clarifies everyone's current (mis)conceptualizations and knowledge. In addition, the teacher encourages students to think of strategies to help move the real image to match the ideal image. In this way, students are creating new knowledge, questioning prior knowledge, and reconstructing knowledge based upon their experience throughout the strategy.

As these three DBI strategies are intended to illustrate, the techniques used in DBI generally fall into three types of instruction as described by Chi: active, constructive and interactive, though no DBI activity is likely to be *only* one type of strategy at all times. That is, all DBI strategies support multiple ways for participants to actively engage in learning. This classification of DBI strategies as typically active, constructive, and/or interactive is important in light of the extensive research suggesting that there is a hierarchy among the strategies in terms of their effectiveness. Namely, research suggests that interactive strategies support learning outcomes to a greater extent than constructive strategies, constructive strategies support learning better than active strategies, and active strategies support learning outcomes to a greater extent than passive strategies. With this in mind, it is expected that DBI strategies will enhance student learning outcomes compared to the traditional passive classroom strategies.

Experiential Learning Cycle

In the previous sections, I discussed drama-based instructional strategies as a collection of individual strategies. In addition to being a set of discrete instructional strategies, DBI is an approach to pedagogical practice within the educational setting that is informed by theories of instruction and learning. A DBI lesson or a series of lessons is carefully constructed to include multiple DBI strategies that as a whole reflect progressive teaching models and learning in a student-centered environment. As such, in thinking about why we might expect DBI to effectively enhance learning and learning-related outcomes of the student, it is helpful to consider the tenets of the learning theories on which it is based.

Based on the work of progressive educators and researchers (Dewey, Lewin, and Piaget), Kolb developed the Experiential Learning Cycle (ELC) as a way to develop instruction that couples hands-on learning with curricular goals and objectives (Kolb, 1984). He posited that learners need to have concrete experiences, reflect on their observations, develop an abstract conceptualization of the experience, and finally try out their ideas through active experimentation. In sum, the experience serves as a source for reflection, observation, and then action in order for the learner to have a deeper, more complete understanding of the concept.

Overall, experiential learning cycle is considered by educational researchers as a theoretical model with strong face validity. Further, various studies have shown that students who learn content through an experiential learning cycle have enhanced learning outcomes (Abdulwahed & Nagy, 2009; Herz & Merz, 1998; Montgomery & Brown, 1997; Nagda, Gurin, & Lopez, 2003) and an increase in self-regulatory behaviors (de Jong, 2006; Kaul & Pratt, 2010). For example, Herz and Merz (1998) found that students who experienced simulation games based on the ELC model had significantly greater gains in content knowledge in economics than students who learned through lecture only.

In terms of the application of the experiential learning cycle to DBI, as would be theoretically recommended, DBI learners first participate in a *concrete experience*. For example, participants could be asked to participate in a town hall meeting where the students are in role as community members who need to make a significant decision for their community. After this concrete experience, the learners move into *reflective observation* to consider the shared concrete experience. This reflection could take many different forms. For example, students may be asked to write a journal entry as their character or create a series of frozen images as a way to respond to the town hall meeting. At this point in the DBI lesson, the teacher may share specific information about the

problem that is facing the community. This may be the point in the lesson where a teacher draws out the larger environmental or community impacts of their decisions. This is a form of *abstract conceptualization*. Students have had an experience to reflect upon but now the teacher directs their attention to relevant abstract concepts that will inform the final step in the cycle: *active experimentation*. During this step, students use what they have learned in the previous three steps, create a hypothesis about what will happen, and then "try out" or experiment with their ideas. This is yet another opportunity for the students to go in role and see what happens. For example, students could improvise a scene between community members who hold opposing views.

Overall, specific DBI strategies follow this same pattern. For example, a leader facilitates the participants in a theatre game called *Data Processing.* When given a directive, students line up by their first names alphabetically without speaking. Next the facilitator asks the students to reflect on how well they achieved the goal of lining up. In addition, the facilitator asks what strategies the students used that worked or did not work. Then the participants discuss what strategies they would want to use the next time. Finally, the facilitator gives another directive: line up by birth date without talking. This time the students use their hypothesized strategies to better complete the task.

In sum, DBI techniques are designed to move through the ELC in the theoretically prescribed manner: A) first, engaging in concrete experimentation through experiencing a town hall meeting or playing the game the first time, B) next, reflecting on the concrete experience through discussion with the facilitator, C) then hypothesizing about what strategies or ideas the students should try next, and finally, D) experimenting with their ideas by improvising a scene or playing the game again. Since DBI maps directly onto the ELC, it is expected that students who learn content through DBI will demonstrate

greater academic achievement compared to students who learn through traditional methods.

Social Constructivism

Drama-Based Instruction is also situated in pedagogical beliefs and practices informed by social constructivism (Vygotsky, 1978). Three prominent ideas from social constructivism serve as the theoretical underpinnings of DBI, including: social and cultural understanding of the learners, scaffolding the learning through the learner's zone of proximal development, and co-constructing meaning through dialectical interactions with others and the environment.

When ascribing to a social constructivist perspective, a leader approaches learning by acknowledging that everyone enters the learning situation with prior knowledge and experiences that can support or inhibit learning. This is in direct opposition to the "banking method" of teaching that assumes that the leader is the expert and that students are "empty vessels" to be filled with knowledge (Freire, 2007). A teacher in a constructivist environment acknowledges and incorporates participants' experiences. For example, rather than giving students a worksheet that outlines the advantages and disadvantages to technology; a facilitator might invite participants to brainstorm pros and cons to using school funding toward technology support. This incorporates students' ideas into the learning while also preparing them to engage in a persuasive discussion supporting either side of the argument.

Second, the facilitator scaffolds the learning such that each student is able to participate at an appropriate level of challenge. For example, in *Post-It Dialogue*, a

facilitator may invite students to respond on individual post-it notes to statements or questions about the curricular topic. Then small groups of students categorize their responses to reflect what the class knows about the topic. Then, each group shares with the classmates and facilitator their categories. Finally, the facilitator can invite students to create frozen pictures to represent and embody these ideas. In this process, students who do not have as much to contribute (due to lack of experience or knowledge in the area) can share information anonymously on the post-its. Then in the small groups, these students will have their ideas confirmed or adjusted as they fit with other group members' ideas. In addition, the facilitator can gauge the level of understanding of the class and adjust her teaching accordingly with a type of dynamic assessment (Poehner, 2007). In sum, learners offer a wealth of knowledge and experiences that the facilitator and/or expert others incorporate into the lesson through scaffolding the learning to strengthen academic and social outcomes (Shamir, 2005; Shamir & Lazerovitz, 2007).

As the final component relevant to DBI, students learn in a social constructivist classroom by co-constructing meaning through dialectical interactions between the learner, the facilitator, and the task/environment (Wells, 2007). "Learners are influenced by, and at the same time push back, take from, change, control, and create an environment in which learning is situated" (Alexander, Schallert, & Hare, 1991, p. 180). To illustrate this idea, I reference the example in the previous paragraph. When students write their responses on the post-it notes, they work to retrieve information from their long-term memory. Once they are in the small groups, the students respond to the comments on the post-its in relationship with their own comments. In response to the

facilitator's prompt, they then create categories for their comments. It is likely that if this were an individual activity, the resulting categories and knowledge would be different. In sum, the learners co-construct knowledge about the world and their experiences by interacting with classmates, the facilitator, and the environment.

The somewhat limited research on social constructivist teaching suggests that students learning through a social constructivist strategies leads to greater social and academic gains (Barnett et al., 2008; Terwel, Oers, Dijk, & Eeden, 2009), deeper connections (Webb & Mastergeorge, 2003), and students make more unique contributions in class (Yager & Akcay, 2008). For example, Barnett et al. (2008) found that preschoolers who learned language and math curriculum through a social constructivist based curriculum had significantly greater gains in social development and less problem behavior than students who learned in a traditional classroom.

DBI explicitly attempts to align whole lessons and individual strategies with social constructivist theory. For example, during the strategy, *Image Work,* the leader invites participants to create a frozen image with their bodies of a concept, e.g., freedom or justice. Then the class interprets the frozen image through a process of describe, analyze, and relate. Students describe the details of the image that are based on concrete evidence, e.g., her hand is placed over her heart. Then the leader asks the students to analyze what that image might mean, e.g., she is saying the pledge or she loves her country. Finally, the leader asks the students to relate the image to an example of freedom that they may or may not see in their lives. Throughout this process, the leader invites the social and cultural understanding of the learner to influence their interpretations of the

image. In addition, the leader is scaffolding their learning through building on familiar ideas to relate to larger concepts. Finally, the leader encourages multiple interpretations of the images such that the students' understandings of the concept may be broadened, challenged and/or confirmed.

In sum, to the extent that DBI strategies use techniques that are in line with recommendations of social constructivist theory, it is expected that students in DBI classrooms will have a deeper understanding of content material and exhibit more appropriate behaviors compared to students in traditional classrooms.

Self-Determination Theory

Finally, we might also expect DBI to lead to benefits in learning and learning-related outcomes in consideration of motivation theories. These theories suggest that the way instruction is structured can have meaningful consequences for students' motivation, engagement, and eventual academic achievement (e.g. Ryan & Deci, 2000). Specifically, Self-Determination Theory (SDT) posits that environments that support basic psychological needs for autonomy, competence, and relatedness to a greater extent will be more facilitative of students' engagement and persistence (Deci & Ryan, 1985, 2008a, 2008b).

Research has supported the benefits of supporting these psychological needs in the classroom. Students' psychological need satisfaction in the classroom has been found to positively relate with their interest in the academic material (Minnaert, Boekaerts, & De Brabander, 2007), academic engagement (Park, Holloway, Arendtsz, Bempechat, &

Li, 2012) and flow or total involvement in a task (Kowal & Fortier, 1999). In the following, I discuss each psychological need in more detail.

The need for autonomy reflects an individual's desire for her or his actions to emanate from the self (DeCharms, 1968; Ryan & Deci, 2000). Autonomy-supportive environments have been shown to support adaptive educational outcomes, such as lower drop-out rates (Alivernini & Lucidi, 2011), adaptive goals toward learning outcomes (Vansteenkiste et al., 2010) and student engagement (Hyungshim, Reeve, & Deci, 2010). In practice, teachers may be autonomy-supportive by providing choices, acknowledging student affect, and giving opportunities for self-direction (Ryan & Deci, 2000). Whereas, teacher practices that potentially diminish autonomy are offering no choice, using controlling language, and imposing strict deadlines (Ryan & Deci, 2000).

The need for competence is the desire of individuals to be effective in interactions with the environment (Harter, 1982; Ryan & Deci, 2000). Conditions that support competence possess an optimal challenge, structure, and a tolerance of failure. If a task is perceived by the student to be too hard, then they may not engage in the behavior; whereas, if the task matches the student's level of ability and effort then the student will persist (Ryan & Deci, 2000). Research suggests that perceived competence correlates with positive academic achievement and mastery goals (Cho, Weinstein, & Wicker, 2011), academic motivation (Faye & Sharpe, 2008), and positive affect and engagement (Miserandino, 1996). Teacher practices that facilitate competence are offering optimal challenges, providing effectance-promoting feedback, and avoiding demeaning feedback (Ryan & Deci, 2000). Whereas, teacher practices that undermine competence in students

are offering tasks that are too easy or too hard and negative performance feedback (Ryan & Deci, 2000).

The need for relatedness (or belongingness) reflects most individuals' desire to establish a close emotional bond and attachment with other people (Baumeister & Leary, 1995; Ryan & Deci, 2000). People have a desire for social interactions and want relationships where they perceive to be understood, accepted, valued, and cared for by others. When this need is met, people are more resilient to stressful situations, seem to have better academic outcomes (Beachboard, Beachboard, Li, & Adkison, 2011), have stronger academic engagement (Furrer & Skinner, 2003), and more academic persistence (Hausmann, Schofield, & Woods, 2007). Although this part of SDT has not been studied as much, a teaching practice that seems to support this need is providing a warm and caring environment. Whereas, teacher practices that may not support relatedness are not connecting with students, facilitating high risk tasks without support for students, and showing negative affect (Ryan & Deci, 2000)

In line with theory and research from self-determination theory, many practices of drama-based instruction would also seem supportive of psychological needs. DBI may be highly supportive of autonomy to the extent that it tends to makes extensive use of student decision-making and encourages students to work in their own way. More specifically, standard in DBI is for the leader to set up a structure for learning (e.g., provide instructions or introduce the material to be learned) and then allow students to direct the action and make choices about learning throughout the process. Although this kind of process is present in many types of instructional strategies, DBI tends to

capitalize on the use of student choice by incorporating student ideas into current and future iterations of the DBI strategy. For example, the leader may ask the students to reflect on their work during a theatre game, rate their performance, decide on strategies to make it better, and then try the theatre game again. "The learner is seen as having the same agentive footing in the interaction as the teacher. . . the student is seen as active, influencing the teacher while being influenced" (Hicks, 1996, p. 30).

Likewise, consistent with the tenets of SDT, the structure of DBI supports a student's need for perceived competence. In DBI, a facilitator scaffolds the learning by varying the level of vulnerability of the learner. For example, within a role drama, students may create characters based upon their prior knowledge. If they have little prior knowledge of the specific topic, they are able to participate to the extent that they are comfortable. In addition, a DBI teacher supports failure tolerance as many of the strategies are built upon the idea that students may begin to understand the why and how of a concept if they are given the opportunity to experiment with their ideas multiple times. In sum, scaffolding the learning and failure tolerance support the student's need for perceived competence in the learning task.

Finally, DBI practices may be supportive of relatedness by design. Many of the strategies cull on the collection of theatre games that in their very nature were intended to facilitate community development and team-building among the participants. For example, in *Three Ball Toss*, the teacher begins with a group challenge to throw the balls with the intention that the participant can catch the ball. This is emphasized by providing positive practice of what a good catch and throw can look like. For another example, DBI

strategies incorporate students' prior experiences and knowledge. This allows for the opportunity to connect to other students and create a sense of belongingness.

In sum, the design of DBI is such that it is expected to readily support the psychological needs of students (autonomy, competence, and relatedness). In particular, effectively implementing DBI strategies in the classroom means that leaders are intentionally attempting to offer meaningful choices, give opportunities for experimenting with ideas, scaffold learning, support failure tolerance, develop community, and create a sense of belongingness. Given this enhanced support for psychological needs, it is expected from a self-determination perspective that DBI will better support students' engagement in tasks, persistence toward goals, and academic achievement compared to the traditional classroom.

Potential Negative Effects of DBI

While the proposed effects of DBI have overwhelmingly been assumed to be positive, researchers have also suggested potential negative effects of DBI on academic and academic-related outcomes. In particular, the most prevalent and persistent critiques of DBI are 1) the inability for a teacher to cover a breadth of material and 2) the student's cognitive overload (Sweller, 1994) may prevent or inhibit engagement with the new content (Eisner, 1998; R. Smith, 1995; Winner & Cooper, 2000). Both of these possible situations may result in less learning and negative student affect.

Typically, using a DBI strategy to teach curriculum may take more time. Students need to first learn *how* to learn in this new way and then teachers can introduce the new content. Students may feel overwhelmed by learning new content in a physically and/or

cognitively demanding method (Perkins, 1991). In other words, the teacher in DBI asks the students to not only learn X content in a roundabout way but *also* to learn that material X in a new way. From the students' perspective, they may believe that they will learn better and more efficiently if they were given the information in a straight forward manner. This may reduce positive affect and attitude toward the classroom content and learning method (Perkins, 1991).

PAST RESEARCH ASSESSING THE EFFECTS OF DRAMA-BASED INSTRUCTION AND ITS LIMITATIONS

Although educators have led forms of drama-based instruction since the turn of the century, formal research studies were first conducted and documented in the late 1960's. Initially, studies focused on two broad areas of interest for elementary students: language, such as written and oral communication, and personal development, such as self-concept, creativity, and pro-social behaviors, including attitudes toward children with disabilities. Starting in the 1980's, research assessing the effects of DBI broadened to a wider range of student samples, including secondary and post-secondary students, and began to include a broader range of outcomes across academic areas including science and math.

When looking at research over the last twenty-five years, the outcomes measured in DBI research can be categorized into three categories: 1) immediate and delayed achievement and learning in non-drama curriculum, 2) immediate and delayed achievement and learning in drama curriculum, and 3) other psychological and social outcomes related to learning, including engagement, attitudes toward school and

academics, academic self-concept, self-regulation, school attendance, and pro-social attitudes and behavior. Table 3 presents a list of outcomes that have been the focus of DBI research and examples of how these outcomes have been measured.

Outcome	Specific Measures
Immediate/long-term achievement and learning in non-drama curriculum	Standardized tests in various domains (i.e. math, language arts, science, etc…) Unit tests/Chapter tests in various domains Portfolio assessments in various domains
Immediate achievement and learning in drama curriculum	Observation reports of drama skills Self-reports of drama knowledge
Other psychological and social outcomes potentially related to learning	Observation reports of engagement Self-reports of attitudes toward school and academics Researcher developed measures of indicators of psychological health Incidence of problem behavior reports Attendance records Self-reports of pro-social attitudes Incidence of pro-social behavior reports

Table 3. DBI Outcomes and Measures.

In the following section, I consider DBI outcomes and relevant research from each of the categories outlined in Table 3.

Achievement and learning in non-drama curriculum

A great deal of research has examined the effects of DBI on immediate achievement across a variety of subject areas. However, despite the many hypothesized learning benefits of DBI, findings have been mixed. For example, while some studies on the effect of DBI with language arts curriculum have suggested that DBI leads to positive effects on oral and written language outcomes (Hendrickson & Gallegos, 1972; Moore & Caldwell, 1990; Niedermeyer & Oliver, 1973; Wagner, 1986, 1990), other studies have found neutral or negative effects of DBI on oral and written language outcomes (Harris & Rosenberg, 1983; Ingersoll & Kase, 1970; Lawton & Brandon, 2005; Stewig & Vail, 1985). For example, Moore and Caldwell (1990) found that students who participated for 15 weeks in 45-minute drama sessions that focused on developing characters had significantly greater gains in their writing skills than students who received the instruction in language arts during the same period. Alternatively, Lawton and Brandon (2005) conducted research on the effects of DBI on reading achievement as measured by a national standardized test. At the end of the first year of implementation, no significant differences were found between the treatment and control schools when controlling for pre-test assessment of the outcomes.

In contrast to the mixed results in language arts, the effects of DBI on achievement outcomes in other domains have been predominantly positive. For example, many studies have found positive DBI effects on academic achievement in science

(Braund, 1999; Dorion, 2009; Francis, 2007; Kase-Polisini & Spector, 1994; Sloman & Thompson, 2010; Warner & Andersen, 2004), foreign language acquisition (Bournot-Trites, Belliveau, Spiliotopoulos, & Seror, 2007; Erdman, 1991; Shacker & et al., 1993), and math achievement (Fleming, Merrell, & Tymms, 2004; Kayhan, 2009; Walker, Tabone, & Weltsek, 2011) among others.

Though not as prevalent as other achievement areas, outcomes measuring 21st century skills also present a mixed picture. These studies focus on one of the designated 21st century skills such as creativity, collaboration, or communication. For example in one study in a theatre class, the leader focused on communication skills over the course of 18 weeks (Ballou, 2000). At the end of the intervention, students actually displayed negative gains in communication skills per an observer rating scale. However, another study focused on creativity in the language arts classroom for 20 weeks found significant positive gains in creativity (Fischer, 1989).

Achievement and learning in drama curriculum

Though less prevalent, some studies have measured outcomes in drama learning (e.g., role-taking ability). Many of these studies have found positive results (Huntsman, 1982; Karakelle, 2009; Rosenberg & et al., 1983; Wright, 2006; Yeh & Li, 2008) but some have found no effect for DBI on affective responses to characters (Harris & Rosenberg, 1983). For example, Wright (2006) found that upper elementary students who participated in 10 weeks of creative drama sessions experienced significant growth in role-taking abilities as compared to control students. In contrast, Harris and Rosenberg

(1983) conducted a study of sophomore English students and their ability to display and discuss an affective response to literature from a character's point of view. Even though the creative drama was specifically focused on role taking and affective responses, the researchers found no significant difference between the treatment and control groups after 15 weeks of instruction.

Psychological and social outcomes potentially related to learning

In addition to studying academic gains, some studies have focused on the effects of DBI on psychological and social outcomes potentially related to learning, including attitudes toward academics and school, self-perception competencies, and pro-social attitudes and behavior. No clear picture has emerged regarding the effects of DBI on school and academic attitudes. That is, while some studies have found DBI to enhance attitudes (Bournot-Trites, et al., 2007; Fleming, et al., 2004; Francis, 2007; Gourgey, Bosseau, & Delgado, 1984; Walsh-Bowers & Basso, 1999), other studies have found no difference in the effects of DBI on attitudes toward academics (Freeman, Sullivan, & Fulton, 2003; Kariuki & Humphrey, 2006). For example, Bournot et. al. (2007) assessed the attitudes of upper elementary students learning in a French immersion class through DBI strategies. The researchers found a significant difference in attitudes toward learning French for students in the experimental group as compared to students who learned in a more traditional way. However, Kariuki and Humphrey (2006) found that students in a classroom where DBI was used to teach math had significant gains in achievement but no change in attitude toward math when compared to a control group. In a similar study,

Fleming et. al. (2004) found that students in the control group actually had significantly more positive attitudes toward math than those in the experimental group.

The effect of DBI on self-perception outcomes (i.e., self-concept, self-discrepancy, and self-efficacy) is similarly mixed. For example, while one study (Fleming, et al., 2004) found that students who participated in DBI in their math classrooms had a significantly more positive self-concept in math than a matched control group of students, other studies have suggested that DBI has no effect on self-worth (Huntsman, 1982).

A number of studies have revealed positive effects of DBI on student pro-social attitudes and behavior, including attitudes toward marginalized groups such as older adults (Bramwell, 1990), Lesbian/Gay/Bisexual/Transgender youth (Aldredge, 2010; Hanley & Gay, 2002; Zanitsch, 2009), victims of bullying (Burton, 2010; Merrell, 2005), and students with disabilities (Miller & Rynders, 1993). For example, Merrell (2005) found that high school students who participated in a DBI bullying prevention program for seven weeks were more likely to take an active role if they witnessed bullying than students who did not participate in the program.

In sum, a range of outcomes have been studied in an attempt to understand the effects of DBI. In all, research findings have largely revealed inconsistent findings across most categories of outcomes, ensuring that the debate that started in the 1960s regarding the overall effectiveness of DBI has continued to the present day.

PREVIOUS REVIEWS OF DBI INTERVENTION RESEARCH

As suggested earlier, DBI may have a critical, positive impact on instruction and learning in the classroom. To that end, researchers and practitioners need to have a nuanced understanding of how DBI may be most profitably used for educational outcomes. To this end, four meta-analytic studies have been conducted to synthesize the effects of DBI on academic and social/emotional outcomes: Kardash and Wright (1986), Conrad (1992), Conrad and Asher (2000), and Podlozny (2000).

Twenty-five years ago, Kardash and Wright conducted the first research synthesis and meta-analysis of two decades of research on the effects of drama-based instruction on the reading skills, communication skills, person perception, and drama skills among elementary students (1986). Additional meta-analyses have been conducted examining the effects of drama-based instruction on academic outcomes (Conrad, 1992), on self-concept and self-esteem (Conrad & Asher, 2000), and on oral and written language skills (Podlozny, 2000).

Results from early research syntheses of the effects of DBI on language arts outcomes (Kardash & Wright, 1986; Conrad, 1992) have presented a mixed picture. While results from the earliest meta-analysis conducted by Kardash and Wright (1986) suggested that drama-based strategies had a positive effect on oral language skills (d = .46), but a negative effect for reading (d = -.05) among elementary aged students. Other meta-analyses conducted by Conrad (1992) and Podlozny (2000) suggested the effects of DBI were unequivocally positive across various language arts outcomes.

More specifically, synthesizing research conducted on DBI with elementary and secondary students through 1990, Conrad's meta-analysis found positive effects of DBI on academic outcomes (1992) including achievement in reading (d = .24), vocabulary (d = .29), oral language (d = .50), and writing (d = .77). Similarly, Podlozny (2000) found significant positive effects on seven categories of language skills including: oral measures of story understanding (r = .24), written measures of story understanding (r = .50), reading achievement (r = .20), reading readiness (r = .25), oral language development (r = .30), vocabulary (r = .06), and writing (r = .29) among preschool through third grade students.

Beyond these achievement outcomes, the only early meta-analysis to have examined the effects of DBI on academic outcomes in domains other than language arts found a positive effect on math achievement (d = .29) (Conrad, 1992). An early research synthesis of the effects of DBI on drama skills only among elementary school students suggested that that the effect was positive, but varied in magnitude depending on the particular outcome (d = .24 to d = 1.53) (Kardash & Wright, 1986).

Mixed findings have also emerged across early meta-analyses of DBI effects on psychological outcomes. While one early meta-analysis found that drama-based strategies had a positive effect on self-esteem among elementary students (d = .42) (Kardash & Wright, 1986), another found no effect of DBI on self-concept and/or self-esteem among elementary students (d = .01) (Conrad & Asher, 2000). Although they used studies from a very similar timeframe, this finding is contradictory to the earlier Kardash and Wright report. Relatedly, one meta-analysis of the effects of DBI among elementary school

students suggested that it had a positive effect on moral reasoning (d = .61) (Kardash & Wright, 1986).

Interestingly, these meta-analyses found multiple significant moderators. The average effect estimates were stronger for gifted as opposed to remedial reading students (Kardash & Wright, 1986; Conrad, 1992). Stronger effects were present for younger as opposed to older elementary students and female versus male students (Kardash & Wright, 1986). This moderator is similar to findings in Conrad's meta-analysis (1992) who found stronger effect sizes for younger populations (preschool) rather than older populations (elementary and secondary); however, Conrad found no effect for gender across the studies. Longer duration of the treatment and greater experience level of the teacher also seemed to have a more positive impact on the effects sizes (Kardash & Wright, 1986). Additionally, Conrad found a smaller effect when the DBI was led by a Drama specialist, presumably with more arts experience, than a classroom teacher (1992). Overall, all the researchers reported that many studies did not include enough information to determine effect sizes; therefore, some of the subgroup effects were calculated based on as few as two studies.

These previous meta-analyses offered research that was both informative and foundational to future studies; however, much contemporary research does not reference these seminal pieces. Prominent researchers in the field have eschewed the current state of affairs in drama-based educational research. In particular, they have suggested that current research is not responsive to previous studies and is not well grounded in educational or instructional theory (Wagner, 1998). Part of this may be due to the fact

that many reviews exclusively feature studies that report positive results as the only evidence for practice. “Although there is much material published that claims the arts cause academic achievement scores to increase . . . it is often difficult to know the basis upon which the claims are made” (Eisner, 1998, p. 52). Researchers continue to call for research that incorporates mixed-methodologies and is in conversation with previous findings (Eisner, 1998; Fleming, et al., 2004; Wagner, 1998).

One effort to begin the process of building on previous findings was *Critical Links: Learning in the arts and student academic and social development* (Deasy, 2002). It offers a review of the research in drama-based instruction (and other arts disciplines) to recommend lines of research based upon “strong” studies that assess the effects of DBI on educational outcomes. This report is very informative and reviews the “best” of over 500 research reports in arts integration; however, it is difficult to ascertain what the cumulative research suggests and the implications for practice. Of the 19 studies that measure the effects of DBI, the authors only offer broad claims and recommendations for practice that are not based on the selected studies. Based on this study, many authors and policy-makers have made very large, unfounded claims about the effects of arts integration. In response, many arts integration researchers have called for more cautious directives about the effects of arts integration in the classroom and the need for nuanced perspectives on the effects of DBI (Eisner, 1998; Fleming, et al., 2004; Mages, 2008; Wagner, 1998; Winner & Cooper, 2000).

FACTORS THAT MAY INFLUENCE THE EFFECTS OF DBI

Mixed findings suggest that the effects of DBI may be very complex and not unequivocally beneficial. It would seem likely that DBI may be more or less effective for particular outcomes and under certain conditions. In particular, characteristics of the sample, characteristics of the intervention, and the nature of the outcome may all influence the observed effectiveness of DBI.

Research on DBI interventions suggests that the effects of DBI on academic outcomes may vary depending on the age of the participants (Podlozny, 2000). Two meta-analyses examining the effect of DBI on academic outcomes with elementary students suggested that the effect size was inversely related to the age of the student through elementary school (Conrad, 1992; Kardash & Wright, 1986). In line with this finding, as students advance to middle and upper elementary, they may begin to focus less on play and more on productive "work" (Erikson, 1959). Thus, DBI may be seen as simple play and not as beneficial for older students.

However, in contrast, Podlozny (2000) found in her meta-analysis among elementary students that while the participants' age was a significant moderator, DBI had a more positive effect on older elementary students. As students reach middle elementary age, they have moved away from egocentric thinking and into more concrete and logical understanding (Piaget, 1952). Thus, this finding might be explained by the early focus on the self that could supplant many of the key attributes of DBI, namely belongingness and role playing. Consequently, DBI may have a larger effect on older elementary students. It is unclear whether DBI will have a differing effect on secondary students as well. Only

one early meta-analysis included secondary students and has been able to compare the effects of DBI across all school age students (Conrad, 1992). However, in that synthesis there was no difference between elementary (d = .43) and secondary students (d = .39), though the effect of DBI on language arts skills and achievement among these groups was smaller than the effect on preschoolers (d = .85). In sum, we can expect that the effects of DBI may be moderated by the developmental maturation of the participants.

The achievement level of the students may also influence the effects of DBI on learning and learning-related outcomes. As discussed earlier, social constructivist environments may overwhelm learners who have little prior knowledge in the targeted area of instruction (Perkins, 1991). Thus, for struggling learners, DBI may have null or even detrimental effects. In line with this notion, early meta-analyses found that DBI had a stronger positive effect on the language arts skills of typically developing or gifted students compared to students who were considered remedial (Conrad, 1992).

Alternatively, it seems possible that the increase in autonomy, competence, and relatedness as supported by DBI may result in greater gains by lower achieving students who are most in need of motivational support. No single research to date has explicitly compared the effects of DBI on outcomes for high achieving students versus low achieving students of similar ability levels, making it difficult to predict a pattern of effect. However, DBI interventions have been conducted with student populations of various achievement backgrounds. As such, an updated meta-analysis may provide an

opportunity to examine whether the effects of DBI vary as a function of achievement level.

DBI may be differentially effective for female compared to male students but it is unclear why this difference may exist. The existing research has been inconsistent regarding the role of gender. While some studies report stronger effects of DBI for girls in oral language skills (Laurin, 2010) and social skills training (Walsh-Bowers & Basso, 1999); in direct contradiction to this finding, other studies have suggested that there is no difference in effects of DBI by gender (Conrad, 1992; Freeman, et al., 2003). One meta-analysis found a slight advantage of DBI on language arts academic outcomes for groups that had more participating females than males in the research study (Kardash & Wright, 1986), while another meta-analysis found no gender difference (Conrad, 1992). In general, language development occurs earlier and language performance is better among females compared to males, even among children as young as 2 or 3 (Burman, Bitan, & Booth, 2008). This may explain why DBI may have a stronger effect for females in language-based outcomes. However, the role of gender remains uncertain in the effects of DBI on academic outcomes in other domains.

A number of characteristics of the intervention, including the duration of the intervention, the type of drama-based strategy (i.e., active vs. interactive), and the training and implementation of DBI by the facilitator may also influence the effectiveness of DBI.

One factor that may contribute to the effectiveness of DBI is the duration of the intervention. If students are familiar with learning in a more traditional classroom setting, then it may take time to adjust to more active strategies and to be an active participant. Logistically speaking, even students who are readily active in learning will need practice in how to participate in theatre games or role-playing in an educational setting. Based on past meta-analyses, there seems to be little difference in interventions that last between a few days to a few weeks but there may be a significantly stronger effect for more intensive DBI interventions that span twelve weeks to a year or more (Conrad, 1992; Deasy, 2002). In contrast, one meta-analysis found no significant difference in effects by the length of the intervention (Conrad & Asher, 2000). Complicating matters more, Kardash and Wright found that the effect size was inversely related to the minutes per session, positively related to sessions per week, and no relationship to the span of the DBI treatment (Kardash & Wright, 1986). In sum, while we might expect that interventions that include brief individual sessions frequently over a long period of time may yield stronger findings, mixed findings in past DBI research suggests that this moderator deserves further investigation.

Another characteristic, the type of DBI intervention may also have differential effects on the outcome. As noted earlier, research suggests that interactive strategies have significantly better effect on learning outcomes as compared to constructive strategies. In addition, constructive strategies have a positive, stronger effect on learning outcomes as compared to active strategies. In Kardash and Wright's meta-analysis (1986), the researchers found that the type of creative drama treatment was a significant moderator.

Although they did not classify the strategies using the same instructional strategy framework (Chi, 2009), it seems that they allude to two similar categories. Treatments that incorporated improvisation strategies effected student outcomes by half of a standard deviation greater than treatments that only incorporated story dramatization (Kardash & Wright, 1986). Typically, improvisation requires students to build upon prior knowledge through dialogue with another person (i.e., interactive); whereas, story dramatization usually involves retelling a story through movement and repetition of a story (i.e., constructive). Thus this suggests that the interactive strategy had a more positive effect than the constructive strategy.

The effects of DBI may also vary depending on the experience level of the facilitator and how the lesson is developed and/or delivered. With only a basic understanding of DBI, facilitators may lead a series of theatre games which will encourage students to be active (Chi, 2009), but may not use constructive and interactive instruction with the students. Thus, we might expect to find DBI to have stronger positive effects when delivered by more experienced facilitators. Prior research has provided support for this hypothesis. For example, Stewig and Vail (1985a) replicated a previous study that found strong effects of DBI on oral language outcomes when a highly experienced teacher led the session. In their follow-up study, they used a leader who had no experience in DBI and found no significant effect of DBI on the oral language outcomes. They suggested that the explanation for the differing effects might be the experience level of the leader. Likewise, Kardash and Wright found in their meta-analysis

that the largest effect sizes were associated with studies that were led by the most experienced facilitators (Kardash & Wright, 1986).

Two characteristics of the outcome may also contribute to the effect of DBI including: the alignment between the outcome and the treatment and the targeted domain for the DBI intervention. The alignment between the measured outcome and the DBI treatment may influence the observed effect. For example, a study that measures standardized math achievement as an outcome of DBI use in the Language Arts curriculum is poorly aligned (Walker, et al., 2011). Alternatively, a study may use DBI to teach role-taking ability and then measure a student's ability to empathize and assume alternative perspectives (Wagner, 1986). This would be considered a close or proximal alignment between the treatment and the outcomes measured. It would be expected that more proximal or closely aligned treatment and outcomes would show a higher effect size than poorly aligned treatment and outcomes.

Thus far, no meta-analysis has tested the outcome domain of interest as a potential moderator for the effects of DBI. For example, does DBI focused on social studies have a greater effect on academic achievement in that domain than DBI focused on math skills? It may be that DBI is more closely related to language arts concepts and skills rather than mathematics concepts and skills. This relationship may facilitate a stronger, more positive effect in domains that are thought to be related to DBI, e.g., language arts, speech, social studies and less effect in domains that are seemingly not as closely related to DBI, e.g., mathematics, science. In order to contribute to a fuller understanding of the impact of DBI, this potential moderator needs to be explored.

THE PRESENT META-ANALYSIS

Despite the undeniable contributions of early syntheses (Kardash & Wright, 1986; Conrad, 1992; Conrad & Asher, 2000; Podlozny, 2000), an updated meta-analysis of drama-based instruction is warranted. There are a number of limitations of the previous meta-analyses and research synthesis. Namely, previous reviews are at this point outdated and limited in scope in terms of the outcomes assessed. These reviews focus solely on language arts, self-concept, or focus mainly on elementary age students. Twenty-five years of research has accumulated since the most comprehensive meta-analysis assessed multiple drama, non-drama and psychosocial outcomes (Kardash & Wright, 1986). A cursory review of the most recent research addressing the effects of drama-based instruction makes it apparent that the outcomes of focus in DBI research has become significantly broader in the last twenty-five years. That is, recent research has not only focused on the effects of DBI on verbal achievement and self-concept, among the other outcomes of focus in early syntheses, but has also focused on the effects of DBI on additional outcomes such as creativity, attitudes toward school, pro-social attitudes and behavior, as well as achievement in science and math. Finally, three of the earlier meta-analyses focused on elementary age students and one of the meta-analyses included participants up to eighth grade. A cursory look at the literature suggests that recent research on effects of DBI have also been conducted with middle school, high school, and college students. As such, this project will summarize research examining the effects of DBI among pre-school through college students.

This project aims to synthesize the research using meta-analysis to address the following research questions:

1) What does the cumulative research suggest regarding the impact of drama-based instructional strategies on student outcomes including academic outcomes, psychosocial outcomes, and 21st century skills?
2) Do characteristics of the intervention, students, or outcomes influence the magnitude or direction of the effect of drama-based instructional strategies?

In response to these research questions, I predict the following:

1. Overall, the effect of drama-based instruction across all student outcomes will be positive.
2. There will be stronger, positive effects for lower elementary age students compared to older students.
3. There will be a stronger, positive effect for female students compared to male students.
4. At this point, the research literature seems inconclusive about how DBI strategies might affect students of differing proficiency levels. However, theoretically, this would seem to be an important moderator to assess. Likewise, because previous meta-analyses reported effects depending on this moderator it seems important to examine it in this meta-analysis as well.
5. There will be a stronger positive effect for interventions that include frequent, brief sessions that occur over a long period of time compared to interventions in which the sessions are infrequent or the intervention as a whole is brief.

6. There will be a stronger positive effect for interventions that include more interactive strategies compared to constructive strategies. In addition, there will be a stronger positive effect for interventions that include more constructive strategies as compared to active strategies.
7. There will be a stronger positive effect for interventions delivered by more experienced facilitators.
8. There will be stronger positive effects of DBI when the outcome is well-aligned with and more proximal to the intervention.
9. Thus far, the research literature does not present a pattern upon which to suggest a directional hypothesis for the differing effects of DBI on academic domains. I will conduct an exploratory analysis to test academic domain of the outcome as a potential moderator.

Chapter Three Methods

Research syntheses primarily focus on empirical studies and seek to summarize past research by drawing overall conclusions from multiple, separate investigations that address related or identical topics. This project employed state-of-the-art methods to perform the research syntheses (Cooper, Hedges, & Valentine, 2009). These methods involved an approach that views research synthesis as a data gathering exercise and applies criteria similar to those employed to judge the validity of primary research (Cooper, 1998). The approach required (a) precise problem definition, (b) exhaustive and unbiased gathering of the research evidence, (c) careful examination of the strengths and weaknesses of the included research, (d) appropriate methods for data integration, including meta-analysis, (e) cautious interpretation of the cumulative evidence, and (f) complete reporting of the syntheses' methods and results.

Study Inclusion Criteria and Search Strategies

To be included in this research synthesis, several criteria had to be met. Most importantly, each study had to have assessed in some way the relationship between drama-based instruction as defined earlier and a student outcome, including academic achievement in drama or non-drama outcomes, attitudes toward academics or another measure of student psychosocial functioning, such as self-concept, self-esteem, or attitudes toward marginalized groups. The studies included in the meta-analysis must all be experiments or quasi-experiments with at least one experimental and one control group. This means that participants in the intervention condition received a DBI treatment while participants in a control condition did not and instead typically received

traditional instruction. Only studies conducted in educational settings during school hours with preschool through college students were included. Samples may include or be tailored to students who are typically developing or students who have various behavior, emotional, or psychological disabilities. In addition, only studies with samples from the United States, Canada, United Kingdom and Australia were included due to the similarity in typical schooling and shared language. Finally, enough information had to be provided in order to calculate an effect size. For example, excluded studies used samples for countries outside the inclusion criteria (Stinson & Freebody, 2006; Saracayir, 2010; Kayhan, 2009). Other excluded studies did not have a control condition (Rosen, 1987) or reported insufficient information to calculate an effect size (Stewig, 1985). Another main reason for exclusion was the intervention was either not DBI (Carson, 1991; Millin, 1986) or used multiple artforms (Brandon, 2004; Karafelis, 1986). Studies included in the meta-analysis met all the criteria listed above (Ballou, 2000; Fleming, 2004; Walker, 2011).

Using a broad set of search strategies, an attempt was made to identify and retrieve the entire population of published and unpublished studies that examine the relationship between drama-based instruction and student outcomes since the last major review of the literature (Kardash & Wright, 1986). First, I searched the *PsychInfo, ERIC, Dissertation Abstracts International, Academic Search Complete, the International Bibliography of Theatre and Dance, American Economic Association* and *Google Scholar* electronic databases for documents catalogued since October 1985 (the end time

point that Kardash and Wright used to gather reports in their meta-analysis) using each of the following keywords:

Creative Drama*, Drama Based, Applied Theatre, Process Drama, Arts Integration

These search terms captured a wide base of studies including studies that did not meet the inclusion criteria, e.g., multi-art, production focused, etc. Creative drama is the most common term used for DBI; however, researchers and practitioners may also refer to DBI as creative dramatics or dramatization. Drama-based captures any research referencing drama-based instruction and drama-based pedagogy. Applied theatre captures a wide range of studies that frequently report outcomes in areas other than achievement and outside of school contexts. Process drama is a common term used in the U.K. to encapsulate DBI research. Finally, arts integration generally references multi-art programs, but these reports were read to confirm this assumption.

In addition, *Social Sciences Citation Index* database was searched for documents catalogued since 1985 that had been cited by previous meta-analyses. Next, I employed three strategies to directly contact researchers who may have studied drama-based instructional strategies. First, I contacted the dean, associate dean, or chair of the 98 colleges, schools, or departments of education at doctoral-granting institutions of higher education with high research productivity and request that they ask their faculty to share with us any research they have conducted that relates to drama-based instruction. Second, I contacted researchers who have been the first author on two or more articles on drama-based instruction during the past ten years. Third, I contacted the director of research in

the regional labs for educational research. Finally, the reference sections of relevant documents were examined to determine if any cited works might be relevant to our topic. A preliminary search using these strategies located a total of 2,892 non-duplicate, potentially relevant documents. Further inspection of these documents suggested that there were 45 reports that met inclusion criteria. Studies were excluded for various reasons. Most commonly, studies were excluded because they suggested a quasi-experimental design in the language in the abstract; however, upon further investigation, the report did not use these methods. Some studies were excluded due to the country of the sample—most often Turkey and Singapore. Studies were also excluded for lack of statistical information, i.e., no sample size reported.

INFORMATION RETRIEVED FROM PRIMARY RESEARCH

Numerous different characteristics of each study were included in the database. These characteristics encompass six broad distinctions among studies: (a) the research report; (b) the research design; (c) the intervention/drama-based instruction variable; (d) the sample of students; (e) the student outcome measure, and (f) the estimate of the relationship between drama-based instruction and the student outcome. Table 4 outlines the characteristics of studies which we coded and the Appendix provides the complete coding guide. Among the characteristics that we recorded were (a) the number of students, classrooms, schools, districts included in the study as a whole or in the drama-based instruction and traditional instruction conditions at the beginning and end of the study, (b) the duration of the intervention, (c) the experience level of the leader, (d) the

type of DBI strategy used (e) the grade level of the students, (f) gender of students, (g) whether any achievement label was applied to the sample of students (h) timing of the outcome measurement relative to implementation of the intervention, (i) the type and nature of the student outcome measure, and (j) the direction and magnitude of the relationship between drama-based instruction and a student outcome.

Report	Author Year of publication Type of publication
Study information	Study number Type of organization Sampling procedure Research funding
Characteristics of the Drama-Based Instruction Intervention	Minutes per day Days Days per month Weeks Total lessons Total hours Word to describe DBI Theoretical frame Types of DBI Linked to standards and domain Leader of DBI and experience level Type of training Leader characteristics Prescribed intervention Domain of DBI Measurement of integrity
Characteristics of control condition	Business as usual OR record all the relevant information
Setting characteristics	State/country Type of community Type of school
Research design	Research design Attrition Characteristics of matching Local event/contaminant
Sample characteristics	Sample/subsample

Table 4. Cont. next page

	Labels for sample SES Gender Race/ethnicity Age of students Education level
Outcome measured	Type of outcome and subject Respondent Referent How measured Administered
Effect size information	Pretest effect/ unadjusted post-test effect/adjusted post-test effect Control variables When measured Unit of assignment and unit of statistical analysis Covariates for equating Direction of effect size Effect size information Normal distribution Variances roughly equivalent across groups

Table 4. Outline summary of coding guide

Variability in study design. Generally, the studies did not vary in design such that most studies conducted the intervention and then took post measures. No studies randomly assigned the condition at the student level; however, a few studies randomly assigned at the school or classroom level. These latter studies used the student as the unit of analysis and are not considered randomly assigned for the current analyses. Whenever reported in the studies, I gathered information for a pre-test or covariate (e.g., prior achievement) that was used to control for differences in samples before the intervention. In order to retain as much information as possible, I computed effects based on post

effect estimates, but also, computed effects based on the adjusted effect estimates. Occasionally, studies measured outcomes throughout the study (every week); however, so few studies did this that the weekly measures could not be included in the analysis.

Variability in study sample and intervention. Generally, the studies reported a relatively even distribution of females to males. This sample of studies did vary in duration of the intervention, grade level of the intervention, domain of the intervention, and the method for measuring outcomes. For example, the duration of the intervention ranged from one lesson to 100 lessons. Studies used various grade levels as their samples from preschool through college students including: three preschool, 11 lower elementary, 12 upper elementary, 17 middle school, three high school, and two college studies. Although English language arts represented the majority of the studies, many other domains were represented, including: science, social studies, math, theatre, and foreign language. The outcomes for achievement and 21st century skills were measured by standardized tests, tests, observer ratings (e.g., coding a writing sample), and interviews; whereas, the outcome for attitudes toward academics was measured by a survey or a test. Outcomes for self-competency skills were measured by standardized tests, tests, and surveys. Outcomes for social skills were measured by observer ratings (e.g., teachers coding for friendship making), standardized tests, or surveys. Outcomes for motivation were measured by observer ratings, standardized tests, and school record. Attitudes towards others were measured by a test or a survey. Finally, arts outcomes were measured by observer ratings. It is important to note that most of the dissertations in this meta-analysis were housed in a school of education rather than a school of theatre. It is

likely that the methodological beliefs and practices of these two types of schools may differ.

Variability in outcome. It is important to note the variability of studies within each outcome as well. For example, self-perception competency skills included measures of self-concept, self-esteem, self-efficacy, and self-discrepancy. I grouped this set of outcomes together to reflect the students' beliefs about themselves and their ability. Although these outcomes each have specific research literature, many of the terms overlap and are closely related and are categorized as such for this present study. For the outcome of attitudes towards others, studies included measures of attitudes toward elderly, students with disabilities, students as friends, and attitudes toward bullying and bullies. Motivation was measured in various ways including: desire to learn, engagement in the process, absenteeism, time on task, attention, and interest. Outcomes in social skills varied, including: problem behavior, ability to make friends, conflict resolution, cooperation, and recognizing emotion in others. Outcomes related to 21st century skills measured such things as critical thinking, creativity, and self-regulatory skills. Attitudes towards academics included attitudes toward school, peers, and specific academic domains. Finally achievement outcomes included various measures of specific academic content (unit test on French language) as well as more general types of academic content (standardized test of basic skills).

Coder reliability. Two coders extracted information from all reports selected for inclusion. Discrepancies were noted and discussed by the coders and if agreement was not reached, a third coder was consulted.

To this end, I conducted coder training to ensure a comprehensive understanding of each of the codes. For example, I provided an overview of the definition of drama-based instruction as well as a detailed description of each type of strategy that might be described in the research reports. Initially, the coders worked with coding practice documents, coded these documents independently, and then we compared the codes. Based on these sessions, I revised the coding guide. Only after coders demonstrated adequate agreement in training did they independently complete coding on the study documents. Prior to resolving conflicts, the coders had a 94% rate of agreement for 33,000 codes; however, the code for strategies used in the DBI intervention was particularly problematic in that it accounted for 22.7% of the discrepancies.

STATISTICAL PROCEDURES

Effect size estimation. The standardized mean difference or the *d*-index (Cohen, 1988) was used to estimate effects . Calculating the *d*-index for any comparison involves dividing the difference between the two group means by either their average standard deviation or by the standard deviation of the control group. In this synthesis, I subtracted the control condition mean from the DBI intervention condition mean and divide the difference by their average standard deviation. Thus, positive effect sizes indicated that students who received DBI had higher learning and learning-related outcomes than students who did not receive DBI. When available, I calculated effect sizes based on the means and standard deviations of the student outcomes. If means and standard deviations were not available, I retrieved the information needed to calculate *d*-indexes indirectly

from inferential statistics (see Borenstein, 2009; Lipsey & Wilson, 2000). Effect sizes that adjust or control for the outcome variable prior to intervention were also retrieved or calculated if the needed information was available. Adjusted effect sizes were calculated if the needed information was available in the following order of preference:

1. By calculating pre-test and post-test effect sizes separately (based on pre-test M/SD and post M/SD) and taking the difference
2. By using adjusted Ms/SDs but calculating as regular post-test ES
3. By using F-test from ANCOVA but treating it as a regular F test to calculate imprecise adjusted effect size
4. If the post-test scores are mean gain scores (MGSs), then calculate ES as if it is a regular post ES with Ms/SDs.

Methods of data integration. First, the distribution of effect sizes was examined to determine if any are statistical outliers. Grubbs (1950) test, also called "the maximum normed residual test" was applied (see also, Barnett & Lewis, 1994). This test identifies outliers in univariate distributions and does so one observation at a time. If outliers were identified, (using $p < .05$, two-tailed, as the significance level) these values were set at the value of their next nearest neighbor.

Although every attempt was made to obtain all relevant research reports, it is possible that I may not find some reports. To account for this possibility, I used the Duval and Tweedie's (2000a, 2000b) trim and fill procedure to test whether the distribution of effect sizes used in the analyses were consistent with variation in effect sizes that would be predicted if the estimates were normally distributed. If the data was found to be skewed

due to missing reports or biased reporting, then this method filled in an estimated value for the "missing" data to achieve a normal distribution. Finally, the procedure imputed the missing values and offers an estimate of the impact of the aforementioned data censoring on the observed distribution of effect sizes.

I used weighted procedures to calculate average effect sizes across all comparisons (Borenstein, Hedges, Higgins, & Rothstein, 2005). Each independent effect size was first multiplied by the inverse of its variance. The sum of these products was then divided by the sum of the inverses. Also, 95% confidence intervals were calculated for average effects. If the confidence interval did not contain zero, then the null hypothesis of DBI v. non-DBI difference was rejected.

Shifting unit of analyses. For this analysis, I used a shifting unit of analyses (Cooper, 1998). In this procedure, each effect size associated with one study is first coded as if it were an independent estimate of the relationship between DBI and the outcome. For example, if a single sample permitted comparisons of the effect of DBI for both achievement in math and achievement in reading, two separate effects were calculated. However, when estimating the overall effect of DBI, I averaged these effects prior to analysis so that the one sample only contributed one effect size. For the overall average weighted mean and CI, I weighted the effect by the inverse of its variance which is based primarily on sample size. The averaged weighted effect then required that the sample size be about equal for achievement in reading and achievement in math. Then, I meta-analyzed the effect of DBI on all achievement outcomes. However, in an analysis that tested the effects of DBI on reading and math separately, the sample contributed one

effect size to reading and one effect size to math. Then, I conducted separate meta-analyses to test the effect of DBI on outcomes related to math achievement and outcomes related to reading achievement. This method retains as much data as possible from each study while holding to a minimum any violations of the assumption of independent data points.

Possible moderators (e.g. grade level, gender, duration of the treatment, etc.) of the drama-based instruction to student outcome relationship were tested using homogeneity analyses (Cooper, et al., 2009; Hedges & Olkin, 1985). The analyses was carried out to determine whether (a) the variance in a group of individual effect sizes varies more than predicted by sampling error and/or (b) multiple groups of average effect sizes varies more than predicted by sampling error.

Fixed error models assume sampling error is due solely to differences among participants in the study. However, it is also possible to view studies as containing other random influences. If we believe that random variation in DBI interventions was a significant component of error, a random error model should be used that takes into account this study-level variance in effect sizes (see Hedges & Vevea, 1998, for a discussion of fixed and random effects). Thus, rather than opt for a single model of error, I conducted all analyses twice, once employing fixed-error assumptions and once employing random-error assumptions. This sensitivity analysis allows me to examine the effects of the different assumptions (fixed or random) on the findings. Differences in the results may influence interpretation of the results. For example, if a moderator is found to be significant under a fixed effect assumption but not significant under a random effect assumption, then this

suggests a limit on the generalizeability of the inferences of the moderator.

All statistical processes were conducted using the Comprehensive Meta-Analysis software package (Borenstein, Hedges, Higgins, & Rothstein, 2005). Only outcome measures that were reported in four or more separate reports with four or more independent samples were meta-analyzed.

Chapter Four Results

The literature search uncovered 45 studies that tested the effect of drama-based instruction on various student-related outcomes. The 45 studies reported 275 separate effect sizes based on 61 independent samples. The authors, sample sizes, effect sizes, and other important study characteristics are listed in Table 5. In this sample of studies, 109 effect sizes were reported for the effect of drama-based instruction on an achievement outcome. All other outcomes included a total of 2 to 12 overall effects collapsed across subgroups. Drama skills had the fewest overall effects, then attitude toward others (2), motivation (6), social skills (6), attitudes toward academics (8), 21st century skills (9), and self-perception competencies (12).

Author (Year)	*Doc type*[1]	*Sample*	*Gen*[2]	*Grade level*[3]	*No. of lessons*	*Lead*[4]	*Word to dscrbe*[5]	*Prox*[6]	*Domain*	*Otcm*[7]	*Outcome type*	*Meas type*[8]	*Effect size*
Arieli (2007)	D	64	NR	Mid	15	R	CD	D	Science	Ach	Recall/Interpretation	ST	1.84(2.00)
											Conceptual	I	1.05
Arise (2008)	O	356	NR	U El	27.5	TA	D	I	General	Att	Toward academics	Sv	-.07(.03)
		357									Toward school		-.09(-.09)
		340								Self	Self-efficacy		-.19(.01)
		24								Att	Toward academics		.06(.13)
		24											-.19(-.12)
		24								21st	Critical thinking		.07(-.03)
Ballou (2000)	D	22	NR	Mid	18	TA	DBI	D	Theatre	Att	Toward school	Sv	1.21(1.83)
										Self	Self-concept		2.02(3.78)
									Theatre	21St	Communication	O	-.90(-.81)
									Other	Mot	Absenteeism	SR	1.00(.56)
Bournot-Trite (2007)	J	60	.58F	Mid	7	CT	DBI	D	Foreign language	Ach	Writing	O	.45
								I		Att	Toward academics	Sv	1.00
										Mot	Interest	Survey	.65
Bramwell (1992)	J	46	.49F	U El	16	CT	D	D	General	Att O	Toward elderly	T	.48
													2.49
													.52
													1.60
Buege (1994)	J	47	NR	U El	32	CT/R	CD	D	Other	Att O	Toward disabilities	Sv	(.26)
Byerly (1994)	J	17	NR	HS	1	NR	D	D	Reading	Ach	Writing	O	-.25
				HS							Writing		.31
Cooney (1999)	D	34	.647F	Coll	15	R	PD	D	Theatre	Arts	Speech	O	.52
											Presence		.55
											Listening		.58
											Concentration		.50
											Clarity		.68
											Energy		.57
Cormack (2003)	T	54	.462F	Mid	20	R	CD	D	LA	Ach	Writing	O	.96(.53)
											Writing	O	.93(-.00)
Danner (2003)	D	42	.77F	Mid	18	R	CD	D	Theatre	Self	Self-efficacy	Sv	.20(-.18)
		46								Att	Toward school		.58(-.25)
		43									Toward academics		-.08(-.34)
		46									Toward parents		-.03(-.60)
		48								21st	Critical thinking	T	-.08(-.08)
Dupont(1992)	J	34	NR	U El	30	R	CD	D	Reading	Ach	Comprehension	ST	1.35(1.14)

Table 5. Cont.

													3.60(2.91)
Enciso	UP	64	.50F	U El	NR	CT/T	DIE	D	LA	Att	Toward academics	Sv	.51(.15)
		33		U El		A		I		Self	Self-efficacy		-.65(-.52)
		61		U El				I		Self	Self-concept		.23(-.05)
		83		U El				I		Att	Toward academics		.15(-.13)
		265		Mid				D					.50(.40)
		114		Mid				I		Self	Self-efficacy		.05(.04)
		218		Mid				I		Self	Self-concept		.05(.07)
		265		Mid				I		Att	Toward academics		-.03(.07)
		83		HS				D					.22(.27)
		81		HS				D					.27(.18)
		81		HS				I			Toward school		.04(.12)
		39		HS				I		Self	Self-concept		-.17(-.15)
Fernsler (2003)	T	30	NR	L El	6	R	D	D	Soc. St.	Ach	Conceptual	T	.77
Fischer (1989)	D	14	NR	Mid	20	R/CT	PD	I	LA	21St	Critical thinking	T	1.55(1.37)
		14											2.24
		26											1.62(1.71)
		21											2.27
		11(high)											5.04
		6(high)											5.03
		14									Creativity		2.92(2.14)
		14											4.98
		23											2.75(5.57)
		21											2.76
		15(high)											1.22(1.10)
		11(high)											1.97
		6(high)											2.17
Fleming (2004)	J	165	NR	U El	15	TA/C T	D	D	Reading	Ach	Comprehension	ST	(.4)
									Math			ST	(.83)
									LA		Writing	O	(.3)
		103						I	General	Att	Toward school	Sv	(.46)
													(.34)
										Att O	Peers		(.19)
										Self			(.62)
											Self-concept		
Francis (2007)	J	22	1F	Mid	8	CT	DBI	D	Science	Ach	Comprehension	T	.72
		19	0F										.99
Freeman (2003)	J	89	.47F	U El	18	OA/T A	CD	D	Other	Self	Self-concept	Sv	-.33(.20)
		91											-.29(.28)

Table 5. Cont.

		90										-.44(-.30)
		94										.03
		93										.02
		94										.26
		93pre/po							SS	With peers		-.27(.67)
		102post										-.03
		93pre/po								Problem behavior		.17(-.16)
		102post										.07
GoldStein (1985)	D	63	.545F	HS	9	R	CD	D	Other	21st	Critical	T	.82(1.33)
		53											2.14
Jackson (2000)	T	39	NR	Mid	7	R	CD	I	Other	Self	Self-concept	Sv	.10(-.30)
Karafelis (1986)	D	25(low)	.512F	U El	100	CT	D	D	Reading	Ach	Comprehension	T	.71
		29(mid)											.34
		23(high)											.01
Kariuki (2006)	C	26	.54F	U El	5	NR	D	D	Math	Ach	Comprehension	T	-1.29
										Mot	Interest	Sv	0
Laurin (2010)	T	45	.627F	L El	8	R	CD	D	LA	Ach	No. of words	O	.15(.14)
											No. of adjective		0(0)
											No. of difft adj.		-.39(.27)
											No. of adverb		.26(.01)
											No. of difft adv.		.19(.08)
											Structure writing		-.02(0)
										Mot	Engagement		-.50(-.82)
McCambridge (1998)	D	53	.50F	Mid	14	R	PD	D	Mixed	21st	Critical thinking	Sv	(-.27)
													(-.06)
McDonald (1992)	D	32	.53F	PreK	7	R	SD	D	LA	Ach	Verbal	O	.71(1.03)
											Communication		.75(1.18)
											Words/unit		-.82(-.51)
											Mazes		1.60(1.95)
											Words/maze		1.01(.83)
											Complexity		.78(1.10)
McFadden (2010)	D	33	NR	L El	9	TA	T	I	Reading	Ach	Literacy	ST	1.20(.11)
													.45(.28)
McGregor (2001)	D	45	.557F	Coll	16	CT	CD	I	Reading	Ach	Vocab/Comprehension	ST	.28(.16)
									Creativity	21st	Critical thinking		.40(.19)
											Creativity		.71(1.22)
								D		21st	Critical thinking		0
									Study habits				

Table 5. Cont.

McNamee (1985)	J	65	NR	PreK Pre K L El	24	CT	D	D	Reading	Ach	Developing stories	O	-.12(-.76) 1.80(1.16) 1.92(1.90)
Merrell (2004)	D	56	.60F	HS	7	R	DIE	D	Commun ication	Att O	Toward bullying	Sv	.03 .28 .24 -.18(.37)
Miller (1993)	J	12	.54F	U El	12	CT	CD	I	Other	SS	With peers	Sv	.54(.71) .93(-.61)
Moore (1993)	J	41	NR	L El	15	CT	D	D I	LA	Ach Att	Writing Toward academics	O Sv	1.53(1.01) .51
Ranger (1995)	T	50	.50F	U El	NR	NR	CD	D	Reading	Ach	Vocab/Comprehension	O	1.22
Rapaport (1989)	D	71	.422F	L El	24	R	CD	D	LA	Ach	Comprehension No. of words Meaningful words Length No. of mazes Original words Story developing No. of unit % of words % picture unit	ST T	.22(.24) .08 .04(.00) .10(.09) .20(-.04) .15(.00) -.25(.23) -.12(.12) -.47(.20) .18(.49) 1.04(.43) .29(.12) .09(-.02) .25(0) 0(-.36) .50(.33) -.08(.39)
Risemberg (1992)	J	48	.50F	L El	4	R	D	D	LA	Ach	Word fluency Vocab diversity No. action words No. emotions Developing story	O	(1.24) (.99) (1.23) (0) (1.59)
Rose (2000)	J	157 147 147 147	NR	U El	20	TA	DBI	D	Reading	Ach Arts Arts	Comprehension Performance Performance	ST O	.59(.21) .55(0) .53 .60
Saab (1987)	D	42	1F	Mid	NR	CT	CD	D	Math	Ach	Comprehension	ST	.01(.09)

Table 5. Cont.

		40	0F									.67(.36)
			1F									-.23(-.39)
			0F									.47(.06)
			1F						Att	Toward academics	Sv	-1.01(-.34)
			0F									.45(-.25)
		41	1F				I		21st	Creativity		.26(-.25)
		41	0F									-23(-.30)
Smith (2010)		83	.58F	L El	13	TA	DBI	I	Other	21st	Creativity	ST	-.08(.06)
													.21(.17)
										Mot	Engagement		.22(.26)
								D		SS	With peers		.28(-.09)
Wagner	D	28	0F	U El	1	R/CT	Other	D	LA	Ach	Writing	O	.06
(1986)		28	1F	U El									.96
		22	0F	Mid									-.12
		22	1F	Mid									.79
		28	0F	U El							No. persuasive		.40
		28	1F	U El									.87
		22	0F	Mid									-.31
		22	1F	Mid									.56
		28	0F	U El							No. mazes		.85
		28	1F	U El									.73
		22	0F	Mid									.30
		22	1F	Mid									.03
		28	0F	U El									-.35
		28	1F	U El									.47
		22	0F	Mid									-.39
		22	1F	Mid									-.39
Walker-a	J	699	NR	Mid-1	40	CT/T	AI	I	Reading	Ach	Comprehension	ST	.33
(2011)		702		Mid-1		A			Math				.20
		338		Mid-2					Reading				.28
		338		Mid-2					Math				.36
Walker-b	J	203	NR	U El	20	CT	AI	D	LA	Ach	Comprehension	ST	.76
(2011)		146							Soc. St.				.21
		336						I	LA	Mot	Engagement	O	.22
										SS	With peers		0
Walsh-	J	104	.54F	Mid	21	R/CT	CD	I	Other	Att	Toward school	Sv	-.45(.10)
Bowers		104								SS	With peers		.39(.03)
(1992)		104											.24(.00)
		48									Problem behavior		-.37(-.35)

Table 5. Cont.

		48										.78(.72)
		104										-.73)-1.34
		104										-.19(.79)
Warner	J	33	.20F	L El	1	R	PD	D	Science	Ach	Comprehension	O	.88
(2004)		35									Writing		1.18
		35											1.55
		36								Mot	Engagement		.93
Weidner	T	30	NR	L El	1	R	CD	D	Reading	Ach	Recall	I	.34
(1993)													
Wright, P.	J	26	0F	U El	15	NR	DIE	D	Reading	Ach	Developing story	T	.52(.52)
(2006)		26	0F								Picture vocab.		-1.00(-.50)
		24	1F								Developing story		.64(1.60)
		24	1F								Picture vocab.		-1.25(.33)
		26	0F					I	Other	Self	Self-discrepancy	Sv	.37(-.14)
		26	0F								Self-concept		.10(.47)
		24	1F								Self-discrepancy		.95(-.57)
		24	1F								Self-concept		-.60).38)
Wright, E.	G	160	NR	L El	18	TA/R	D	D	Reading	Ach	Picture vocab.	T	.17(.07)
(1986)											Story recall		.01(.30)
													-.21(.07)
													-.07(.08)
													-.17(.01)
													.13(.15)
													.07(-.01)
													-.02(-.10)
													.02(-.20)
													-.11(.31)
													-.25(.10)
													-.05(.33)
													-.27(.10)

Note: Some abbreviations are necessary for formatting. NR = Not reported. [1]Under Document type, D = Dissertation; T = Thesis; J = Journal; UP = Unpublished; C = Conference paper; G = Government report. [2] Under Gender, F = % Female in sample. [3] Unger Grade level, PreK = Pre-Kindergarten, L El = Lower Elementary, U El = Upper Elementary, Mid = Middle School, HS = High School, Col = College; [4]Under Type of Leader, R = Researcher; CT = Classroom Teacher; TA = Teaching Artist; [5]Word to describe the intervention, CD = Creative Drama; D = Drama; DIE = Drama in Education; PD = Process Drama; DBI = Drama-Based Instruction; T = Theatre; AI = Arts Integration; [6]Under Proximity between outcome and measure, D = Direct, I = Indirect; [7]Under Outcome, Ach = Achievement, Att = Attitudes toward academics, Self = Self-competency skills, Att O = Attitudes toward others, 21st = 21st century skills, Mot = Motivation, SS = Social skills, Arts = Arts skills.

Table 5. Characteristics of Quasi-Experimental Studies Included in Meta-Analysis.

The 45 studies appeared between 1985 and December 2012. The sample sizes ranged from 8 to 702. For each set of outcomes, Grubb's test was used to identify outliers within that set of sample sizes and within that set of effect sizes. After Winsorizing outliers, all studies were retained for further analysis. For the set of samples assessing the effects of DBI on achievement, four outliers were detected for the *n*'s (n = 699, 702, 338, 338; all from Walker, et. al., 2011). All were Winsorized to the nearest neighbor (*n* = 203). For the set of effects assessing DBI on achievement, one outlier was detected to the right side of the distribution (*d* = 3.60; Dupont, 1992) and was Winsorized to *d*= 1.93. No outliers for sample size were detected for the adjusted effects of DBI on achievement. One outlier for effect size was detected (*d* = 2.91; Dupont, 1992) and Winsorized to *d* = 2.00.

For the set of samples assessing the effects of DBI on self-perception competencies, attitudes toward academics, and 21st century skills, no outliers for sample sizes or effect sizes were detected. For the set of samples assessing the adjusted effects of DBI on self-perception competencies, no outliers for sample sizes or effect sizes were detected. For the set of samples assessing the adjusted effects of DBI on attitude toward academics, no outliers in sample size was detected, but one outlier was detected in effect size estimates to the right side of the distribution (*d* = 3.78; Ballou, 2000) and was

Winsorized to .47. For the set of samples assessing the adjusted effects of DBI on 21st century skills, no outliers for sample size was detected; however, one outlier for effect sizes was detected (d = 5.57; Fischer, 1989) and was Winsorized to the nearest neighbor to d = 2.14.

For the set of samples assessing the effects of DBI on social skills, an outlier for the sample size was detected (n = 336; Walker, et. al, 2001) and was Winsorized to n = 84. No outliers for effect size estimates were detected. For the set of samples assessing the adjusted effects of DBI on social skills, no outliers for sample size or effect size were detected. For the set of samples assessing the effects of DBI on motivation, an outlier for the sample size was detected (n = 338; Walker, et. al, 2011) and was Winsorized to n = 83. No outliers were detected for the effect size estimates.

Because there were three or fewer samples contributing to the average weighted effects, no Grubb's test was conducted for the sample of studies assessing the effects of DBI on attitudes towards others, drama skills or the adjusted effects of DBI on motivation, attitudes toward others, and drama skills.

OVERALL EFFECTS OF DRAMA-BASED INSTRUCTION

Achievement. First I examined the overall effect of DBI on each of the eight outcomes as shown in Table 6. Of the 101 effect sizes assessing the effect of DBI on

achievement, 75 were in a positive direction and 26 were in a negative direction. The effects ranged from d = -1.29 to d = 1.93 (after Winsorization). The weighted average d was .46 under a fixed-error (FE) model with a 95% CI from .37 to .55. The weighted average d was .53 under a random-error (RE) model with a 95% CI from .33 to .71. Therefore the hypothesis that the effect of DBI on achievement is equal to 0 could be rejected under both fixed- and random-error models. Additionally, the tests of the distribution of effect sizes revealed that I could reject the hypothesis that the effects were estimating the same underlying population value, $Q(38) = 141.13$, $p < .001$.

The trim-and-fill analyses were conducted using both fixed- and random-error models. I performed the analyses looking for possible missing effects on the left side of the distribution, thus reducing the size of the positive average d. Using a fixed-effects model, I found evidence that three effect sizes might have been missing. Imputing these values would change the mean effect to d = .38 (95% CI = .29, .46) under a fixed effects and d = .40 (95% CI = .19, .61) under a random effects model. Using the random-effects model, I found no evidence of missing effect sizes. Even when I tested for possible data censoring, the effects of DBI on achievement was positive and significantly different from zero.

Of the 71 adjusted effect sizes assessing the effect of DBI on achievement, 44 were in the positive direction and 27 were in the negative direction. The effects ranged from d = -.76 to d = 2.00 (after Winsorization). The adjusted weighted average for d was .43 under a FE model with a 95% CI from .32 to .54. The weighted average d was .55 under a RA model with a 95% CI from .27 to .83. Therefore the hypothesis that the adjusted effect of DBI on achievement is equal to 0 could be rejected under both fixed- and random-error models. Additionally, the tests of distribution of adjusted effect sizes revealed that I could reject the hypothesis that the effects were estimating the same underlying population value, $Q(21) = 119.65$, $p < .001$.

The trim-and-fill analyses were conducted using both fixed- and random-error models. I performed the analyses looking for possible missing effects on the left side of the distribution, thus reducing the size of the positive average d. Using a fixed-effects model, I found evidence that four effect sizes might have been missing. Imputing these values would change the mean effect to d = .27 (95% CI = .16, .37) under a fixed effects and d = .31 (95% CI = 0, .62) under a random effects model. Using the random-effects model, I found evidence of three missing effect sizes. Imputing these values would change the mean effect to d = .30 (95% CI = .19, .40) under a fixed effects and d = .36 (95% CI = .05, .67) under a random effects model. Thus, even after assessing the

influence of possible data censoring, the effect of DBI on achievement was positive and significantly different from zero under all models except a fixed model with random effects.

Self-perception competencies. Thirteen of the 20 overall effect sizes assessing the effect of DBI on self-perception competencies were in a positive direction and 7 were in a negative direction. The effects ranged from d = -.60 to d = 2.02 (after Winsorization). The weighted average d was .01 (95% CI = -.11, .12) under a FE model and .09 (95% CI = -.13, .32) under a RE model, $Q(12) = 32.65$, $p < .001$. Therefore the hypothesis that the effect of DBI on self-perception competencies is equal to 0 could not be rejected under fixed- and random-effects models. Trim-and-fill analyses indicated no additional effects were missing under a fixed- or random-effects model.

Eight of the 16 overall adjusted effect sizes assessing the effect of DBI on self-perception competencies were in a positive direction and 8 were in a negative direction. The effects ranged from d = -.57 to d = 1.60 (after Winsorization). The adjusted weighted average d was -.01 (95% CI = -.14, .12) under a FE model and .05 (95% CI = -.24, .33) under a RE model, $Q(12) = 33.31$, $p < .001$. Therefore the hypothesis that the adjusted effect of DBI on self-perception competencies is equal to 0 could not be rejected under

fixed- and random-effects models. Trim-and-fill analyses indicated no additional effects were missing under a fixed- or random-effects model.

21st century skills. Eighteen of the 22 overall effect sizes assessing the effect of DBI on 21st century skills were in a positive direction and 4 were in a negative direction. The effects ranged from d = -.90 to d = 5.04. The weighted average was significant under FE model (d = .29, p < .01, 95% CI = .07, .49;) and nearly significant under the RE model (d = .45, p < .10, 95% CI = -.08, .98; Q(10) = 43.58, p < .001). Therefore the hypothesis that the effect of DBI on 21st century skills is equal to 0 can be rejected under fixed-error model, but not under random-error models. Trim-and-fill analyses indicated no additional effects were missing under a fixed- or random-effects model.

Nine of the 14 overall adjusted effect sizes assessing the effect of DBI on 21st century skills were in a positive direction and 5 were in a negative direction. The adjusted effects ranged from d = -.81 to d = 2.14 (after Winsorization). The adjusted weighted average d was .20 (95% CI = 0, .41) under a FE model and .27 (95% CI = -.17, .72) under a RE model, Q(9) = 35.67, p < .001. Therefore the hypothesis that the effect of DBI on 21st century skills is equal to 0 could not be rejected under fixed- and random-error models. Trim-and-fill analyses indicated no additional effects were missing under a fixed- or random-effects model.

Attitudes toward academics. Thirteen of the 22 overall effect sizes assessing the effect of DBI on attitudes toward academics were in a positive direction and 9 were in a negative direction. The effects ranged from d = -1.01 to d = 1.21. The weighted average d was .21 (95% CI = .08, .34) under a FE model and .16 (95% CI = -.12, .44) under a RE model, $Q(7) = 26.11$, $p < .001$. Therefore the hypothesis that the effect of DBI on attitudes toward academics is equal to 0 could be rejected under fixed-effects but not under a random-effects model. The trim-and-fill analyses were conducted using both fixed- and random-error models. Using a fixed-effects model, I found evidence that one effect size might have been missing. Imputing these values would change the mean effect to d = .23 (95% CI = .11, .34) under a fixed effects and d = .19 (95% CI = -.08, .46) under a random effects model. Using the random-effects model, I found evidence of two missing effect sizes. Imputing these values would change the mean effect to d = .23 (95% CI = .11, .34) under a fixed effects and d = .18 (95% CI = -.09, .45) under a random effects model. That is, the effect of DBI on attitudes toward achievement was positive and significantly different from zero under fixed-effects but not under random-effects models even after testing for possible data censoring.

Eleven of the 19 adjusted overall effect sizes assessing the effect of DBI on attitudes toward academics were in a positive direction and 8 were in a negative

direction. The effects ranged from d = -.59 to d = 1.83. The adjusted weighted average d was .16 (95% CI = .07, .25) under a FE model and .13 (95% CI = -.02, .28) under a RE model, $Q(12) = 27.62$, $p < .001$. Therefore the hypothesis that the effect of DBI on attitudes toward academics is equal to 0 could be rejected under fixed-effects but not under a random-effects model. Trim-and-fill analyses indicated no additional effects were missing under a fixed- or random-effects model.

Motivation. Six of the 7 overall effect sizes assessing the effect of DBI on motivation were in a positive direction and one was in a negative direction. The effects ranged from d = -.50 to d = 1.00. The weighted average d was .32 (95% CI = .10, .54) under a FE model and .37 (95% CI = -.02, .76) under a RE model, $Q(5) = 14.41$, $p < .01$. Therefore the hypothesis that the effect of DBI on motivation is equal to 0 could be rejected under fixed-effects model but could not be rejected under a random-effects model. Under a FE model, trim-and-fill analyses indicated evidence that one effect size might have been missing. Imputing this value would change the mean effect to d = .28 (95% CI = .06, .49) under a fixed effects and d = .29 (95% CI = -.09, .67) under a random effects model. Using the random-effects model, I found evidence of one missing effect size. Imputing this value would change the mean effect to d = .28 (95% CI = .06, .49) under a fixed effects and d = .29 (95% CI = -.09, .67) under a random effects model.

When I tested for possible data censoring, the effects of DBI on motivation remain positive and significant under a fixed-effect model; however, these effects were not significantly different from zero under the random-effects model.

Two of the 3 overall adjusted effect sizes assessing the effect of DBI on motivation were in a positive direction and one was in a negative direction. The effects ranged from d = -.82 to d = .56. The adjusted weighted average d was 0 (95% CI = -.33, .33) under a FE model and -.02 (95% CI = -.81, .77) under a RE model, $Q(2) = 9.85$, $p < .01$. Therefore the hypothesis that the effect of DBI on motivation is equal to 0 could not be rejected under fixed- or random-effects models. With so few effect sizes, trim-and-fill analyses were not conducted on attitudes towards others and drama skills.

Social skills. Nine of the 14 overall effect sizes assessing the effect of DBI on social skills were in a positive direction and five were in a negative direction. The effects ranged from d = -.74 to d = .93. The weighted average d was .06 (95% CI = -.12, .23) under a FE model and .06 (95% CI = -.12, .23) under a RE model, $Q(5) = 3.67$, $p > .05$. Therefore the hypothesis that the effect of DBI on social skills is equal to 0 could not be rejected under fixed- and random-effects models. Trim-and-fill analyses indicated no additional effects were missing under a fixed- or random-effects model.

Six of the 11 overall adjusted effect sizes assessing the effect of DBI on social skills were in a positive direction and five were in a negative direction. The adjusted effects ranged from d = -1.34 to d = .79. The weighted adjusted average d was .02 (95% CI = -.23, .20) under a FE model and was .02 (95% CI = -.23, .20) under a RE model, $Q(4) = .31$, $p > .05$. Therefore the hypothesis that the adjusted effect of DBI on social skills is equal to 0 could not be rejected under fixed- and random-effects models. Trim-and-fill analyses indicated no additional adjusted effects were missing under a fixed- or random-effects model.

Attitudes toward others. Eight of the 9 overall effect sizes assessing the effect of DBI on attitudes toward others were in a positive direction and one was in a negative direction. The effects ranged from d = -.18 to d = 2.16. The weighted average d was .33 (95% CI = .05, .62) under a FE model and .41 (95% CI = -.12, .93) under a RE model, $Q(2) = 6.37$, $p < .05$. Therefore the hypothesis that the effect of DBI on attitudes toward others is equal to 0 could be rejected under fixed-effects but not under a random-effects model.

All six of the adjusted overall effect sizes assessing the effect of DBI on attitudes toward others were in a positive direction. The adjusted effects ranged from d = .26 to d = 1.60 (after Winsorization). The adjusted weighted average d was .56 (95% CI = .23,

.90) under a FE model and .61 (95% CI = .02, 1.19) under a RE model, $Q(2) = 5.97$, $p < .05$. Therefore the hypothesis that the adjusted effect of DBI on attitudes toward others is equal to 0 could be rejected under fixed-effects and under a random-effects model.

Drama skills. Both of the effect sizes assessing the effect of DBI on drama skills were in a positive direction. The effects ranged from $d = .53$ to $d = .60$. The weighted average d was .57 (95% CI = .29, .84) under a FE model and .57 (95% CI = .29, .84) under a RE model, $Q(1) = 0$. $p > .05$. Therefore the hypothesis that the effect of DBI on drama skills is equal to 0 could be rejected under fixed- and random-effects models. With so few effect sizes, trim-and-fill analyses were not conducted on attitudes towards others and drama skills.

		Unadjusted *d*-index					Adjusted *d*-index				
		Fixed		Random				Fixed		Random		
Outcome	*k*	*d*	- /+CI	*d*	-/+CI	*Q*	*K*	*D*	- /+CI	*d*	-/+CI	*Q*
Achievement	38	.46***	.37/.55	.53***	.33/.71	141.13***	22	.43***	.32/.54	.55***	.27/.83	119.65***
Attitudes toward academics	12	.25***	.13/.36	.25*	.01/.49	36.59***	8	.13*	.00/.26	.13	-.09/.35	15.78*
21st century skills	10	.29**	.07/.49	.45†	-.08/.98	43.55***	10	.20*	.00/.41	.27	-.17/.72	35.67***
Self-perception	12	.01	-.11/.12	.09	-.13/.32	32.65***	10	-.01	-.14/.12	.05	-.24/.33	33.31***
Motivation	6	.32**	.10/.54	.37†	-.02/.76	14.41**	3	.00	-.33/.33	-.02	-.81/.77	9.85**
Social skills	6	.06	-.12/.23	.06	-.12/.23	3.67	5	.02	-.23/.20	.02	-.23/20	.31
Attitudes toward others	3	.33**	.05/.62	.41	-.12/.93	6.37*	3	.56**	.23/.90	.61*	.02/1.19	5.97*
Drama skills	2	.57***	.29/.84	.57***	.29/84	0.00						

Note. CI = confidence interval; DBI = drama-based instruction.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Table 6. Overall effects of DBI on academic related outcomes

MODERATOR ANALYSES

Next, moderators of the effect of DBI were assessed for achievement, attitudes toward academics, 21st century skills, self-perception, motivation, and social skills. Since attitude toward others and drama skills were not statistically heterogeneous and had few studies, I did not conduct moderator analyses on these outcomes. Moderator testing was only done if variability was present among the moderators (more than two studies populating two or more levels of the analysis). Otherwise, if only one study represented a level for a moderator, then the findings could be considered unstable and unreliable. Before coding the documents, I planned on conducting the following moderator analyses: grade level, gender, achievement level of the sample, duration of the intervention, type of strategies, experience level of facilitator, proximity between the intervention and the measured outcome, and the academic domain of the intervention. Due to limited reporting in the sample of studies, I was unable to conduct analyses on prior proficiency status and type of strategies. In addition, I could not conduct a moderator analysis on the experience level of the facilitator, but did use the type of leader (e.g., classroom teacher or teaching artist) as a way to account for variability in the leader. At post-hoc, I added the following moderator analyses to check for publication bias and/or measurement bias: publication status of the report and type of measure used for the outcome. Additionally,

when possible, I conducted a moderator analyses on the type of outcome within the broader category to see if any patterns emerged.

Achievement Moderator Analyses

I conducted moderator analyses of the effect of DBI on achievement using six of the moderators of theoretical and methodological interest. Table 7 presents these results.

	Unadjusted *d*-index							Adjusted *d*-index						
		Fixed			Random				Fixed			Random		
Moderator	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b
Grade level				3.09			3.88				4.41			.38
Preschool	3	.68***	.35/1.01		.78	-.43/1.98		3	.30†	-.02/.63		.43	-.87/1.74	
Lower elementary	9	.49***	.30/.67		.72***	.23/1.22		7	.45***	.26/.64		.60*	.09/1.12	
Upper elementary	12	.41***	.24/.57		.31*	-.01/.62		6	.36***	.18/.55		.47†	.10/.85	
Middle	11	.48***	.33/.63		.59***	.31/87		5	.70***	.41/.99		.71**	-.08/1.51	
Post middle	3	.21	-.29/.72		.21	-.29/.72								
Gender				1.74			.36				.28			.25
% female	21	β = -.33	-.82/.16		-.31	-.98/.36		9	.20	-.52/.92		.21	-.60/1.01	
Number of lesson plans				2.63†			2.71†							
1-5 lessons	11	.25*	.01/.50		.21	-.18/.59								
6 or more lessons	24	.47***	.37/.57		.58	.36/.81								
Leader				4.85†			.71				12.81***			5.15+
Classroom teacher	16	.51***	.38/.63		.62***	.34/.90		8	.43***	.23/.63		.46	-.13/1.06	
Researcher	13	.61***	.42/.79		.63***	.33/.93		8	.77***	.55/.98		.90***	.36/1.44	
Teaching artist	3	.29**	.07/.50		.39	-.14/.91		4	.22**	.04/.40		.22*	.01/.44	
Proximity between DBI and outcome				4.00**			2.47				.00			.03
Directly related	31	.54***	.43/.65		.59***	.37/.82		19	.42***	.30/.53		.55***	.25/.86	
Indirectly related	7	.36***	.22/.59		.31**	.03/.58		3	.43*	.03/.83		.49	-.16/1.15	
Domain of outcome				25.23***			15.92**				2.89			.14
Language arts	22	.55***	.43/.66		.59***	.35/.83		7	.43***	.22/.63		.49**	.16/.84	
Reading	8	.31***	.13/.49		.37†	-.01/.75		14	.34***	.32/.85		.47**	.15/.80	
Math	5	.21*	.03/.39		.21†	-.35/.47		3	.59***	.33/.86		.37	-.25/.98	
Science	4	1.15***	.76/1.54		1.15***	.76/1.54								
Social studies	2	.31*	.00/.62		.39	-.12/.91								
Publication status				2.50			2.19				3.51			3.68
Dissertation	14	.42***	.24/.61		.43***	.20/.66		9	.46***	.26/.66		.53*	.08/.98	
Thesis	5	.68***	.40/.96		.67**	.23/1.11		2	.14	-.26/.54		.14	-.26/.54	
Journal	17	.55***	.43/.67		.69***	.39/.98		10	.55***	.39/.71		.73**	.25/1.21	
Measure of outcome				5.46			1.16				2.83			.35
Test	10	.43***	.20/.66		.44**	.11/.77		5	.24*	.01/.46		.38†	-.04/.80	
Interview	2	.69**	.17/1.20		.69*	.00/1.39								
Observer rating	16	.68***	.53/.84		.68***	.34/1.02		10	.41***	.25/.57		.56**	.10/1.02	
Standardized test	11	.48***	.35/.61		.58***	.32/.84		10	.45***	.30/.60		.50**	.15/.86	

Note. CI = confidence interval; DBI = drama-based instruction. †$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$. The *k* for each moderator may not add up to the overall *k* from Table 6. Studies were excluded for moderator analyses due to insufficient reporting or some studies contributed more than one effect when the study had multiple levels of the moderator.

Table 7. Treatment and student moderators of the effects of DBI on achievement outcomes.

Grade level. First, I examined the association between the magnitude of effect sizes and the grade level of the sample participants. Studies were divided into categories by grade level: preschool (k = 3), lower elementary (k = 9), upper elementary (k = 12), middle (k = 11), and high school/college (k = 3). All studies were included in this analysis. The average weighted effect of DBI on achievement did not significantly vary for different grade levels under a FE assumptions, $Q(5) = 3.09$; therefore, no further analysis was conducted. The adjusted average weighted effect of DBI on academics did not significantly vary among preschool (k = 3), lower elementary (k = 7), upper elementary (k = 6), and middle school (k = 5) grade levels under the FE assumptions, $Q(3) = 4.41$; therefore, no further analysis was conducted.

Gender. I conducted a regression analysis on the percent of females in a sample for each study. The β represents the slope for females such that if the slope is significant and positive then this suggests that as females in the study increases then the average effect estimate in the sample of studies increases. Alternatively, if the slope is significant and negative, then this suggests that as females in the study increases then the average effect estimate of the sample of studies decreases. For the 21 samples reporting percent of gender in the sample, the average weighted effect estimate of DBI on achievement was negative, but not significant under FE assumptions ($\beta = -.33$, $p > .05$, $Q(1) = 1.74$). No

further analysis was conducted. For the adjusted average weighted effect of DBI on achievement, percent female was negative, but not significant under FE assumptions (β = -.31, $p > .05$, $Q(1) = .36$). No further analysis was conducted.

Number of lessons. The initial categories for number of lessons were 1-5 lessons ($k = 11$), 6-10 lessons ($k = 5$), 11-20 lessons ($k = 8$), and more than 20 lessons ($k = 5$). Three studies were excluded from this analysis because the number of lessons was not reported (Francis, 2007; Ranger, 1995; Saab, 1987). I tested whether the 11-20 lessons were distinct from samples who received 21 or more lessons. Under both fixed- and random-error assumptions, the average weighted effect of DBI on achievement when 11-20 lessons were implemented (FE: $d = .44$, 95% CI = .29, .59; RE: $d = .57$, 95% CI = .34, .64) was not significantly different from the average weighted effect of DBI on achievement when participants received 21 or more lessons (FE: $d = .49$, 95% CI = .34, .64; RE: $d = .63$, 95% CI = .19, 1.06). In addition, each of these categories for number of lessons was compared to the 6-10 average weighted effect for distinct differences (FE: $d = .49$, 95% CI = .20, .78; RE: $d = .49$, 95% CI = .20, .78) and found to not be significantly different. Therefore, these three categories were collapsed to represent 6 or more lessons. This analysis revealed a nearly significant difference in the effect of DBI depending on whether the participants experienced 1-5 lessons or 6 or more lessons under

either a fixed-error model, $Q(1) = 2.63$, $p < .10$, or a random-error model, $Q(1) = 2.71$, $p < .10$, favoring the effects when 6 or more lessons are provided. The adjusted average weighted effect of DBI on academics could not be conducted for the number of lessons moderator due to insufficient variability in this sample of studies.

Leader. Studies reported one of three types of facilitators/leaders for the DBI instruction: the classroom teacher ($k = 16$), the researcher ($k = 13$), or the teaching artist ($k = 3$). Four studies were excluded from the analysis due to insufficiently reporting the type of leader (Byerly, 1994; Kariuki, 2006; Ranger, 1995; Wright, 2006). The analysis of the average weighted effect of DBI on achievement revealed a nearly significant difference depending on whether the intervention was facilitated by the classroom teacher, researcher or teaching artist under a fixed-error model, $Q(2) = 4.85$, $p < .08$, but not under a random-error model, $Q(2) = .71$, $p > .05$. I then proceeded to conduct pairwise comparisons under fixed-effects assumptions only. The largest effect was for DBI interventions lead by the researcher (FE: $d = .61$, 95% CI = .42, .79). Researcher led interventions were significantly different from interventions lead by a teaching artist (FE: $d = .29$, 95% CI = .07, .50; $Q(1) = 4.70$, $p < .05$), but not significantly different from interventions lead by a classroom teacher (FE: $d = .51$, 95% CI = .38, .63; $Q(1) = .76$, $p > .05$).

The adjusted average weighted effect of DBI on academics did significantly vary among type of leader including: classroom teacher (k = 8), researcher (k = 8), and teaching artist (k = 4) under the FE assumptions, $Q(2) = 12.81$, $p < .001$ and nearly significant under RE assumptions, $Q(2) = 5.15$, $p < .08$. Pairwise comparisons were conducted under fixed-error models. The largest adjusted effect was for DBI interventions lead by the researcher (FE: d = .77, 95% CI = .55, .98). Researcher led interventions were significantly different from interventions lead by a teaching artist (FE: d = .22, 95% CI = .04, .40; $Q(1) = 12.81$, $p < .001$), as well as significantly different from interventions lead by a classroom teacher (FE: d = .43, 95% CI = .23, .63; $Q(1) = 4.16$, $p < .05$). Adjusted effect estimates for interventions lead by the classroom teacher or the teaching artist did not significantly differ.

Proximity between DBI and measured outcome. Studies were coded for the proximal alignment between the DBI intervention and the outcome that was measured. An example of an effect that was directly aligned used DBI to develop a student's ability to use descriptive words in writing and the measured outcome was the number of descriptive words used in a writing sample (Cormack, 2003). An example of an effect that was indirectly aligned used DBI to teach comprehension skills when reading novels and the measured outcome was a math achievement test (Walker, et. al., 2011).

This sample of studies included 31 that were directly aligned and 7 that were indirectly aligned. All studies were included in the analysis. The average weighted effect of DBI on achievement significantly varied under fixed-error assumptions, $Q(1) = 4.00$, $p < .01$, but not under random-error assumptions, $Q(1) = 2.47$, $p > .05$, favoring the directly aligned outcomes. Studies reported adjusted effects for directly aligned outcomes ($k = 19$) and indirectly aligned outcomes ($k = 3$). The adjusted average weighted effect of DBI on academics did not significantly vary for proximal alignment under the FE assumptions, $Q(1) = 0$; therefore no further analysis was conducted.

Domain of DBI intervention. Studies were coded for the domain of the DBI intervention and presented five categories: language arts which focused mainly on writing ($k = 22$), reading which focused mostly on reading comprehension ($k = 8$), math ($k = 5$), science ($k = 4$), and social studies ($k = 2$). The weighted average effect of DBI on achievement significantly varied for different domains under both fixed-error assumptions, $Q(4) = 25.23$, $p < .001$, and under random-error assumptions, $Q(4) = 15.92$, $p < .01$. I proceeded to conduct pairwise comparisons under both models. The largest effect was for science (FE: $d = 1.15$, 95% CI = .76, 1.54, RE: $d = 1.15$, 95% CI = .76, 1.54). Science was significantly different from language arts (FE: $d = .55$, 95% CI = .43, .66, $Q(1) = 8.63$, $p < .01$; RE: $d = .59$, 95% CI = .35, .83, $Q(1) = 5.89$, $p < .05$), for

reading (FE: d = .31, 95% CI = .13, .49, $Q(1)$ = 14.78, p < .001; RE: d = .37, 95% CI = -.01, .75, $Q(1)$ = 8.02, p < .01), for social studies (FE: d = .31, 95% CI = 0, .62, $Q(1)$ = 10.98, p < .001; RE: d = .39, 95% CI = -.12, .91, $Q(1)$ = 5.31, p < .05), and for math (FE: d = .21, 95% CI = .03, .39, $Q(1)$ = 18.48, p < .001; RE: d = .21, 95% CI = -.35, .47, $Q(1)$ = 14.26, p < .001). In addition, weighted average effects for DBI on achievement in language arts are significantly different for reading (FE: $Q(1)$ = 4.75, p < .05) and for math (FE: $Q(1)$ = 9.68, p < .01; RE: $Q(1)$ = 4.66, p < .05). No significant differences were found in average weighted effects between reading, social studies, and math outcomes.

Studies reported adjusted effects of DBI on academics in various domains including: language arts (k = 7), reading (k = 14), and math (k = 3). The adjusted average weighted effect of DBI on academics did not significantly vary by domain of the outcome under the FE assumptions, $Q(1)$ = 2.89; therefore no further analysis was conducted.

Publication status. The data presented three categories for publication status for reports assessing the effects of DBI on achievement outcomes: dissertation (k = 14), thesis (k = 5), and journal (k = 17). A significant amount of variability cannot be attributed to this moderator under the fixed-error assumptions ($Q(2)$ = 2.50, p > .05); therefore, no additional analyses was conducted. For the adjusted weighted effect

estimates, the publication status was dissertation (k = 9), thesis (k = 2), and journal (k = 10). A significant amount of variability cannot be attributed to this moderator under the fixed-error assumptions ($Q(2) = 3.51$, $p > .05$); therefore, no additional analyses was conducted.

Measure of outcome. The data presented four categories for measures of the outcomes assessing the effects of DBI on achievement: test (k = 10), interview (k = 2), observer rating (k = 16), and standardized test (k = 11). A significant amount of variability cannot be attributed to this moderator under the fixed-error assumptions ($Q(3) = 5.46$, $p > .05$); therefore, no additional analyses was conducted. Similar results were produced for adjusted weighted effect estimates of DBI on achievement outcomes. Categories included: test (k = 10), interview (k = 2), observer rating (k = 16), and standardized test (k = 11) and the variability was not significantly accounted for by this moderator under fixed-error assumptions ($Q(3) = 2.83$, $p > .05$).

Type of outcome. Due to insufficient variability, I could not conduct moderator analyses for type of outcome.

Attitudes toward Academics Moderator Analyses

I conducted moderator analyses of the effect of DBI on student attitudes toward academics using six of the moderators of theoretical and methodological interest. Overall,

fewer studies reported outcomes measuring attitude toward academics than achievement. Therefore, I collapsed categorical divisions when it was theoretically relevant and appropriate. Table 8 presents the moderator results.

	Unadjusted *d*-index							Adjusted *d*-index						
		Fixed			Random				Fixed			Random		
Moderator	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b
Grade level				.72			.17				.27			.03
Elementary	4	.32**	.11/.53		.32*	.11/53		2	.09	-.18/.36		.09	-.18/.36	
Middle/High school	8	.21**	.08/.35		.23	-.13/.60		7	.17*	.03/.31		.12	-.18/.42	
Gender				2.03			1.08				.64			.06
% female	15	β = -.51	-1.21/.19		β = -.59	-1.70/.52		16	.16	-.53/.86		.11	-.76/.96	
Number of Lessons				10.98***			8.32**				1.34			.32
11-20 lessons	4	.50***	.21/.80		.57**	.11/1.04		2	.49	-.14/1.12		.70	-1.48/2.87	
21 or more lessons	2	-.31	-.69/.07		-.31	-.69/.07		2	.06	-.33/.44		.06	-.33/.44	
Leader				7.16*			3.31				2.53			1.20
Classroom teacher	7	.26***	.13/.39		.29*	-.02/.55		5	.15†	.01/.28		.13	-.02/.28	
Researcher	2	-.29	-.71/.14		-.23	-.81/.34		2	-.03	-.45/.39		-.04	-.47/.40	
Teaching artist	3	.40**	.09/.71		.54	-.21/1.29		2	.50†	-.03/1.02		.87	-.93/2.66	
Proximity between DBI and outcome				11.34***			.89				.46			.07
Directly related	7	.40***	.27/.54		.35*	.03/.66		7	.17**	.03/.30		.14	-.10/.48	
Indirectly related	8	.10	-.02/.22		.16	-.07/.40		5	.10	-.03/.23		.10	-.09/.17	
Domain of outcome				3.43			.60				5.00†			4.26
Language arts	4	.25***	.12/.39		.25***	.12/.39		3	.18*	.04/.33		.18*	.04/.33	
Math	2	-.13	-.61/.36		-.27	-1.70/1.17		2	-.29	-.77/.19		-.29	-.77/.19	
Publication status				.47			.15				.53			0
Dissertation	4	.21	-.17/.59		.28	-.69/1.24		4	-.01	-.38/.38		.16	-.72/1.03	
Journal	4	.33**	.09/.58		.36	-.22/.94								
Unpublished	3	.24***	.10/.38		.24***	.10/.38		3	.15*	.01/.29		.15	.01/.29	
Type of outcome				5.76†			.55				.25			1.37
Academics	9	.27***	.15/.39		.27**	.05/.49		6	.10	-.03/.23		.10	-.03/.23	
Peers	2	.17	-.44/.78		.17	-.44/.78		3	.02	-.37/.40		.02	-.37/.40	
School	5	.04	-.11/.19		.12	-.22/.45		3	.12	-.05/.30		.42	-.16/1.01	

Note. CI = confidence interval; DBI = drama-based instruction. †$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.
The *k* for each moderator may not add up to the overall *k* from Table 6. Studies were excluded for moderator analyses due to insufficient reporting or some studies contributed more than one effect when the study had multiple levels of the moderator.

Table 8. Treatment and student moderators of the effects of DBI on attitudinal outcomes toward academics.

Grade level. First, I examined the association between the magnitude of effect sizes and the grade level of the sample participants. Studies were divided into categories by grade level: elementary (k = 4) and middle/high school (k = 8). All studies were included in this analysis. The average weighted effect of DBI on attitudes toward academics did not significantly vary for different grade levels under FE assumptions, $Q(1) = .72$; therefore, no further analysis was conducted. The adjusted average weighted effect of DBI on attitudes toward academics did not significantly vary between elementary (k = 2) and middle/high school (k = 7) grade levels under the FE assumptions, $Q(1) = .27$; therefore, no further analysis was conducted.

Gender. I conducted a regression analysis on the percent of females in a sample for each study. For the 15 samples reporting percent of gender in the sample, the average weighted effect estimate of DBI on achievement was negative, but not significant under FE assumptions ($\beta = -.51$, $p > .05$, $Q(1) = 2.03$). No further analysis was conducted. For the adjusted average weighted effect of DBI on achievement, percent female was negative, but not significant under FE assumptions ($\beta = .16$, $p > .05$, $Q(1) = .64$). No further analysis was conducted.

Number of lessons. The reported categories for number of lessons were 11-20 lessons (k = 4), and more than 20 lessons (k = 2). Two studies were excluded from this

analysis because the number of lessons was not reported (Enciso & Lee, NP; Saab, 1987). Under both fixed- and random-error assumptions, the average weighted effect of DBI on attitudes toward academics when 11-20 lessons were implemented (FE: d = .50, 95% CI = .21, .80; RE: d = .57, 95% CI = .11, 1.04) was significantly different from the average weighted effect of DBI on attitudes toward academics when participants received 21 or more lessons (FE: d = -.31, 95% CI = -.69, .07; RE: d = -.31, 95% CI = -.69, .07), FE: $Q(1) = 10.98$, $p < .001$; RE: $Q(1) = 8.32$, $p < .01$, favoring the interventions with 11-20 lessons. The adjusted average weighted effect of DBI on academics did not significantly vary for the number of lessons comparing 11-20 lessons (k = 2) and 21 or more lessons (k = 2) under the FE assumptions, $Q(1) = 1.34$; therefore, no further analysis was conducted.

Leader. Studies reported one of three types of facilitators/leaders for the DBI instruction: the classroom teacher (k = 7), the researcher (k = 2), or the teaching artist (k = 3). All studies were included in this analysis. The analysis of the average weighted effect of DBI on attitudes toward academics revealed a significant difference depending on whether the intervention was facilitated by the classroom teacher, researcher or teaching artist under a fixed-error model, $Q(2) = 7.16$, $p < .05$, but not under a random-error model, $Q(2) = 3.31$, $p > .05$. I then proceeded to conduct pairwise comparisons

under fixed-effects assumptions only. The largest effect was for DBI interventions lead by the classroom teacher (FE: d = .27, 95% CI = .14, .40; RE: d = .28, 95% CI = .02, .55). Classroom teacher led interventions were significantly different from interventions lead by a researcher under FE assumptions, but not under RE assumptions (FE: d = -.29, 95% CI = -.70, .14; RE: d = -.23, 95% CI = -.81, .34; FE: $Q(1) = 6.10$, $p < .01$; RE: $Q(1) = 2.56$, $p < .10$), but not significantly different from interventions lead by a teaching artist (FE: d = .40, 95% CI = .09, .71; RE: d = .40, 95% CI = .09, .71; $Q(1) = .55$). However, the weighted effect estimates of interventions lead by researcher and teaching artist were not significantly different from one another ($Q(1) = .55$).

The adjusted average weighted effect of DBI on attitudes toward academics did not significantly vary among type of leader including: classroom teacher (k = 5), researcher (k = 2), and teaching artist (k = 2) under the FE assumptions, $Q(2) = 2.43$; therefore no further analysis was conducted.

Proximity between DBI and measured outcome. Studies were coded for the proximal alignment between the DBI intervention and the outcome that was measured. This sample of studies included 8 that were directly aligned and 8 that were indirectly aligned. All studies were included in the analysis. The average weighted effect of DBI on attitudes toward academics significantly varied under fixed-error assumptions, $Q(1) =$

10.25, $p < .001$, but not under random-error assumptions, $Q(1) = .94$, $p > .05$, favoring the directly aligned outcomes. The adjusted average weighted effect of DBI on attitudes toward academics significantly varied for proximal alignment between directly aligned ($k = 7$) and indirectly aligned ($k = 5$) under the FE assumptions, $Q(1) = 5.17$, $p < .05$, but not under the RE assumptions, $Q(1) = .81$, $p > .05$, favoring directly aligned outcomes.

Domain of DBI intervention. Studies were coded for the domain of the DBI intervention and presented four categories: language arts ($k = 4$) and math ($k = 2$). Four studies were excluded from the analysis because they reported the domain as general academics or only one study populated a category (Ballou, 2000 reported general communications & Bournot-Trites, 2007 reported foreign language). The weighted average effect of DBI on attitudes toward academics did not significantly vary under fixed-error assumptions, $Q(1) = 2.18$. Therefore, no further analysis was conducted. Studies reported adjusted effects of DBI on attitudes toward academics in various domains including: language arts ($k = 3$) and math ($k = 2$). The adjusted average weighted effect of DBI on attitudes toward academics varied among domains nearly significantly under the FE assumptions, $Q(1) = 2.95$, $p < .09$ as well as under RE assumptions, $Q(1) = 2.95$, $p < .09$; however, no further analysis was conducted.

Publication status. The data presented three categories for publication status for reports assessing the effects of DBI on achievement outcomes: dissertation (k = 4), unpublished (k = 4), and journal (k = 3). A significant amount of variability cannot be attributed to this moderator under the fixed-error assumptions ($Q(2) = .47$, $p > .05$); therefore, no additional analyses was conducted. For the adjusted weighted effect estimates, the publication status was dissertation (k = 4) and unpublished (k = 3). A significant amount of variability cannot be attributed to this moderator under the fixed-error assumptions ($Q(1) = .53$, $p > .05$); therefore, no additional analyses was conducted.

Measure of outcome. I was unable to conduct moderator analyses on the type of measure due to little variability in measurements.

Type of outcome. In this sample of studies, three categories for type of outcome were present, including: attitudes toward academics (k = 9), attitudes toward peers (k = 2), and attitudes toward school (k = 5). This moderator did not account for a significant amount of variability in the outcomes under fixed-error assumptions ($Q(2) = 5.76$, $p > .05$). No further analyses were conducted. For the adjusted average weighted effect estimates, the data presented three categories: attitudes toward academics (k = 6), attitudes toward peers (k = 3), and attitudes toward school (k = 3). This moderator did not

account for a significant amount of variability in the outcomes under fixed-error assumptions ($Q(2) = .25$, $p > .05$). No further analyses were conducted.

Self-Perception Competencies Moderator Analyses

Studies that measured self-perception competencies included outcomes such as self-concept, self-efficacy, and self-discrepancy (for examples, Ballou, 2000; Danner, 2003; Wright, 2006, respectively). I conducted moderator analyses of the effect of DBI on self-perception competencies using five of the moderators of theoretical and methodological interest. Overall, fewer studies reported outcomes measuring self-perception than achievement. Therefore, I collapsed categorical divisions when it was theoretically relevant and appropriate. Table 9 presents the moderator results.

		Unadjusted *d*-index							Adjusted *d*-index						
		Fixed			Random				Fixed			Random			
Moderator	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b	*K*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b
Gender				.01			.01				.26			1.17
% female	7	β = .06	-.99/1.12		β = .06	-1.00/1.12		6	β = -.64	-1.69/.42		β = -.64	-1.69/.42	
Grade level				.74			.80				.26			1.17
Elementary	7	-.03	-.18/12		.02	-.25/.29		5	-.04	-.22/.13		-.04	-.22/.13	
Middle/High school	5	.07	-.12/.26		.26	-.20/.72		5	.03	-.17/.22		.33	-.32/.97	
Leader				1.95			2.51				1.48			1.02
Classroom teacher	3	-.06	-.25/.13		-.06	-.25/.13		3	-.07	-.26/.12		-.08	-.28/.12	
Other arts teacher	2	-.12	-.41/.17		-.12	-.57/.33								
Teaching artist	3	.10	-.10/.30		.68	-.24/1.60		2	.12	-.12/.35		1.83	-1.86/5.53	
Proximity between DBI and outcome				.17			.75				2.48			1.72
Directly related	4	.06	-.21/.32		.37	-.36/1.09		3	.26	-.10/.62		1.08	-.60/2.76	
Indirectly related	8	-.01	-.14/.13		.03	-.18/.24		7	-.05	-.19/.09		-.05	-.19/.09	
Domain of outcome				.25			.56				.20			.23
General academics	7	.00	-.15/.16		.07	-.21/.35		5	-.01	-.20/.18		-.10	-.20/.18	
Language arts	3	-.06	-.25/.13		-.06	-.25/.13		3	-.07	-.26/.12		-.08	-.28/.12	
Publication status				8.92**			2.37				5.88*			1.12
Dissertation	2	.90**	.26/1.54		1.08	-.70/2.86		2	.83*	.12/1.53		1.76	-2.12/5.64	
Journal	5	.15	-.07/.37		.15	-.24/.54		3	.03	-.31/.36		.03	-.31/.36	
Unpublished	3	-.06	-.25/.14		-.06	-.25/.13		3	-.07	-.26/.12		-.08	-.28/.12	
Type of outcome				10.81**			7.56†				.28			.89
Self-discrepancy	2	.64*	.06/1.21		.64*	.06/1.21		2	-.03	-.60/.54		-.05	-1.07/.98	
Self-efficacy	4	-.13	-.30/.03		-.16	-.45/.12		4	-.05	-.21/.12		-.02	-.21/.16	
Self-image	2	-.15	-.44/.15		-.15	-.51/.21								
Self-concept	10	.11	-.02/.24		.14	-.13/.40		8	.01	-.13/.16		.12	-.23/.47	

Note. CI = confidence interval; DBI = drama-based instruction.

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

The *k* for each moderator may not add up to the overall *k* from Table 6. Studies were excluded for moderator analyses due to insufficient reporting or some studies contributed more than one effect when the study had multiple levels of the moderator.

Table 9. Treatment and student moderators of the effects of DBI on self-perception competencies outcomes.

Grade level. First, I examined the association between the magnitude of effect sizes and the grade level of the sample participants. Studies were divided into categories by grade level: elementary (k = 7) and middle/high school (k = 5). All studies were included in this analysis. The average weighted effect of DBI on self-perception competencies did not significantly vary for different grade levels under FE assumptions, $Q(1) = .74$; therefore, no further analysis was conducted. For the adjusted effect estimates, studies divided into elementary (k = 5) and middle/high school (k = 5). The adjusted average weighted effect of DBI on self-perception competencies did not significantly vary for different grade levels under the FE assumptions, $Q(1) = .26$.

Gender. I conducted a regression analysis on the percent of females in a sample for each study. For the 7 samples reporting percent of gender in the sample, the average weighted effect estimate of DBI on achievement was positive, but not significant under FE assumptions ($\beta = .06$, $p > .05$, $Q(1) = .01$). No further analysis was conducted. For the adjusted average weighted effect of DBI on achievement, percent female was negative, but not significant under FE assumptions ($\beta = -.64$, $p > .05$, $Q(1) = 1.17$). No further analysis was conducted.

Leader. Studies reported one of three types of facilitators/leaders for the DBI instruction: the classroom teacher (k = 3), other arts teacher (k = 2), or the teaching artist

(k = 3). One study was excluded from this analysis due to insufficient reporting (Wright, 2006). The analysis of the average weighted effect of DBI on self-perception competencies revealed no significant variation for different types of leaders under FE assumptions, $Q(2) = 1.95$; therefore, no further analysis was conducted. Studies reported adjusted effects included: the classroom teacher (k = 3), and the teaching artist (k = 2). The adjusted average weighted effect of DBI on self-perception competencies did not significantly vary for different types of leaders under the FE assumptions, $Q(1) = 1.48$.

Proximity between DBI and measured outcome. Studies were coded for the proximal alignment between the DBI intervention and the outcome that was measured. This sample of studies included 4 that were directly aligned and 8 that were indirectly aligned. All studies were included in the analysis. The average weighted effect of DBI on self-perception competencies did not significantly vary under fixed-error assumptions, $Q(1) = .17$; therefore, no further analysis was conducted. Studies reported adjusted effects for directly aligned outcomes (k = 3) and indirectly aligned outcomes (k = 7). The adjusted average weighted effect of DBI on self-perception competencies did not significantly vary for proximal alignment under the FE assumptions, $Q(1) = 2.48$.

Domain of DBI intervention. Studies were coded for the domain of the DBI intervention and presented four categories: general academics (k = 7) and language arts

(k = 3). All studies were included in the analysis. The weighted average effect of DBI on self-perception competencies did not significantly vary under fixed-error assumptions, $Q(1) = .25$. Therefore, no further analysis was conducted. Studies reported adjusted effects of DBI on self-perception for general academics (k = 5) and language arts (k = 3). The adjusted average weighted effect of DBI on self-perception competencies did not significantly vary for domain of the outcome under the FE assumptions, $Q(1) = .20$.

Measure of outcome: Due to insufficient variability, I could not run a moderator analyses on the different measures for the outcome.

Publication status. As an additional check for publication bias, I conducted moderator analyses for the publication status. The data presented three categories for publication status for reports assessing the effects of DBI on self-perception competency skills: dissertation (k = 2), unpublished (k = 5), and journal (k = 3). A significant amount of variability can be attributed to this moderator under the fixed-error assumptions ($Q(2) = 8.92$, $p < .01$); therefore, additional analyses were conducted under fixed-error assumptions only. The largest effect of DBI was for dissertations (FE: d = .90, 95% CI = .26, 1.54, $p < .01$). The effects reported in dissertations were significantly different from effects in journals (FE: d = .15, 95% CI = -.07, .37, p>.05; $Q(1) = 4.76$, $p < .05$) as well a significantly different (FE: d = -.06, 95% CI = -.25, .19, p>.05; $Q(1) = 8.03$, $p < .01$).

Effects for journals and unpublished reports were not significantly different from each other (FE: $Q(1) = 2.09$, $p > .05$).

For the adjusted weighted average effect estimate, the data presented three categories: dissertations ($k = 2$), journals ($k = 3$), and unpublished data ($k = 3$). The outcome varied significantly for publication status under fixed-effects models only ($Q = 5.88$, $p > .05$); therefore, additional analyses were conducted under fixed-error assumptions only. The largest adjusted effect of DBI in this sample of studies for dissertations (FE: $d = .83$, 95% CI = .12, 1.53, $p < .05$). Adjusted effects in dissertations significantly differed from adjusted effects from journals (FE: $d = .03$, 95% CI = -.31/.36, $p>.05$; $Q(1) = 4.03$, $p < .05$) as well as from unpublished reports (FE: $d = -.07$, 95% CI = -.26/.12, $p>.05$; $Q(1) = 5.84$, $p < .05$). Adjusted effects from journals and unpublished reports did not significantly differ (FE: $Q(1) = .26$, $p > .05$).

Type of outcome. In this sample of studies, four categories for type of outcome were present, including: self-discrepancy ($k = 2$), self-efficacy ($k = 4$), self-image ($k = 2$), and self-concept ($k = 10$). This moderator accounted for a significant amount of variability in the outcomes under fixed-error assumptions ($Q(3) = 10.81$, $p < .01$) and nearly significant under random-error assumptions $Q(3) = 7.56$, $p = .06$). Further analyses were conducted under fixed-error models only. The largest effect of DBI was on

measures of self-discrepancy (FE: d = .64, 95% CI = .06, 1.21, p < .05). Self-discrepancy was significantly larger than the effects of DBI on self-efficacy (FE: d = -.13, 95% CI = -.30/.03, p>.05; Q(1) = 6.37, p < .01) as well as significantly greater than the effects of DBI on self-image (FE: d = -.15, 95% CI = -.44/.15, p>.05; Q(1) = 5.64, p < .05). Alternatively, the effects of DBI on self-discrepancy were not significantly greater than the effects of DBI on self-concept (FE: d = .11, 95% CI = -.02/.24, p>.05; Q(1) = 3.10, p > .05). Effects of DBI on self-efficacy and self-image were not significantly different (FE: Q(1) = .01, p > .05); however, they are significant compared to self-concept (FE: Q(1) = 5.17, p < .05). Finally, the effects of DBI on self-image and self-concept are not significantly different (FE: Q(1) = 2.40, p > .05).

For the adjusted average weighted effect estimates, the data presented three categories: self-discrepancy (k = 2), self-efficacy (k = 4), and self-concept (k = 8). This moderator did not account for a significant amount of variability in the outcomes under fixed-error assumptions (Q(3) = .28, p > .05). No further analyses were conducted.

21st Century Skills Moderator Analysis

Studies that measured 21st century skills included outcomes such as creativity and critical thinking (for examples, McGregor, 2001; Fischer, 1989, respectively). I conducted moderator analyses of the effect of DBI on 21st century skills using six of the

moderators of theoretical and methodological interest. Overall, fewer studies reported outcomes measuring 21st century skills than achievement. Therefore, I collapsed categorical divisions when it was theoretically relevant and appropriate. Table 10 presents the moderator results.

	Unadjusted *d*-index							Adjusted *d*-index						
		Fixed			Random				Fixed			Random		
Moderator	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b
Grade level				6.55*			2.09							
Elementary	2	.07	-.32/.46		.07	-.32/.46								
Middle	6	.08	-.27/.43		.50	-.39/1.38								
Post Middle	3	.64***	.29/.99		.67	-.17/1.51								
Gender				.76			.76							
% female	2	β = -1.94	-6.33/2.44		β = -1.94	-6.33/2.44								
Number of lesson plans				.52			1.19							
6 to 10 lessons	2	-.03	-.51/.45		-.03	-.51/.45								
11 or more lessons	7	.17	-.10/.45		.42	-.24/1.08								
Leader				7.25*			4.41				11.92**			11.60**
Classroom teacher	3	.13	-.24/.50		.13	-.25/.51		3	.09	-.29/.47		.07	-.60/.73	
Researcher	4	.57***	.23/.90		1.10*	.05/2.16		2	1.70***	.76/2.64		1.70*	.76/2.64	
Teaching artist	3	-.09	-.45/.26		-.19	-.77/.39		3	-.06	-.42/.29		-.16	-.70/.39	
Proximity between DBI and outcome				.02			.78				.90			.00
Directly related	5	.23	-.06/.52		.14	-.55/.84		4	.31*	0/.62		.23	-.60/1.07	
Indirectly related	6	.20	-.09/.50		.59	-.11/1.29		6	.11	-.18/.39		.26	-.28/.80	
Domain of outcome				26.19***			23.01***				3.33			4.08
General academics	4	.43**	.14/.71		.47	-.14/1.08		3	.20	-.13/.53		.20	-.13/.53	
Theatre	2	-.45	-1.05/.14		-.47	-1.28/.33								
Math	2	-.03	-.51/.45		-.03	-.51/.45		2	-.28	-.76/.20		-.28	-.76/.20	
Language arts	2	2.77***	1.61/3.93		2.77***	1.61/3.93		3	.27	-.19/.72		.95	-.51/2.42	
Measure of outcome				5.98*			3.25				28.45***			28.45***
Test	3	.84**	.18/1.50		1.75	-.47/3.97		3	1.43***	.93/1.94		1.43***	.93/1.94	
Standardized test	2	.17	-.18/.52		.17	-.18/.52		2	.24	-.12/.59		.24	-.12/.59	
Survey	4	-.07	-.40/.26		-.07	-.40/.26		4	-.21	-.53/.12		-.21	-.53/.12	
Type of outcome				.58			.46				.84			.57
Creativity	5	.49**	.14/.85		.94*	.03/1.86		5	.49**	.13/.84		.77	-.17/1.70	
Critical thinking	8	.33**	.09/.56		.68*	.00/1.14		6	.29*	.04/.53		.36	-.13/.85	

Note. CI = confidence interval; DBI = drama-based instruction. $^{+}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$. The *k* for each moderator may not add up to the overall *k* from Table 6. Studies were excluded for moderator analyses due to insufficient reporting or some studies contributed more than one effect when the study had multiple levels of the moderator.

Table 10. Treatment and student moderators of the effects of DBI on 21st century outcomes.

Grade level. First, I examined the association between the magnitude of effect sizes and the grade level of the sample participants. Studies were divided into categories by grade level: elementary (k = 2), middle school (k = 6), and high school/college (k = 3). All studies were included in this analysis. The average weighted effect of DBI on 21st century skills significantly varied for different grade levels under FE assumptions, $Q(2) = 6.55$, $p < .01$, but not under RE assumptions, $Q(2) = 2.09$. Pairwise comparisons were conducted under fixed-error models only. The largest effect was for high school/college samples (FE: .64, 95% CI = .29, .99). The weighted average effect of high school/college was significantly different under FE assumptions from the weighted average effect for elementary students (FE: .07, 95% CI = -.32, .46; $Q(1) = 4.65$, $p < .05$) as well as significantly different from the average effect size estimate for middle school students under FE assumptions (FE: .08, 95% CI = -.27, .43; $Q(1) = 4.94$, $p < .05$). However, middle school students effects were not significantly different from elementary school student effects ($Q(1) = 0$. The studies reporting adjusted effects did not have variability in order to conduct analysis on those effect estimates.

Gender. I conducted a regression analysis on the percent of females in a sample for each study. For the 7 samples reporting percent of gender in the sample, the average weighted effect estimate of DBI on 21st century skills was negative, but not significant

under FE assumptions (β = -1.94, $p > .05$, $Q(1) = .76$). No further analysis was conducted. I was unable to conduct analysis on adjusted effect estimates due to limited reporting.

Number of lessons. The reported categories for number of lessons were 6-10 lessons ($k = 2$), and 11 or more lessons ($k = 7$). One study was excluded from this analysis because the number of lessons was not reported (Saab, 1987). Under both fixed- and random-error assumptions, the average weighted effect of DBI on 21st century skills when 6-10 lessons were implemented was not significantly different from the average weighted effect of DBI on 21st century skills when participants received 11 or more lessons, $Q(1) = .52$. Therefore, no further analysis was conducted. The studies reporting adjusted effects did not have variability in order to conduct analysis on those effect estimates.

Leader. Studies reported one of three types of facilitators/leaders for the DBI instruction: the classroom teacher ($k = 3$), the researcher ($k = 5$), or the teaching artist ($k = 3$). Four studies were excluded from the analysis (Arise, 2008; Ballou, 2000; McGregor, 2001; Smith, 2010). The analysis of the average weighted effect of DBI on attitudes toward 21st century skills revealed a significant difference depending on whether the intervention was facilitated by the classroom teacher, researcher or teaching artist

under a fixed-error model, $Q(2) = 7.25$, $p < .05$, but not under a random-error model, $Q(2) = 4.41$, $p < .05$. I then proceeded to conduct pairwise comparisons under fixed-effects assumptions. The largest effect was for DBI interventions lead by the researcher (FE: $d = .57$, 95% CI = .23, .90). Researcher led interventions were significantly different from interventions lead by a teaching artist (FE: $d = -.09$, 95% CI = -.45, .26, $Q(1) = 6.93$, $p < .01$), but nearly significant from interventions lead by a classroom teacher (FE: $d = .13$, 95% CI = -.24, .50, $Q(1) = 2.91$, $p < .08$).

The adjusted average weighted effect of DBI on 21st century skills significantly vary among type of leader including: classroom teacher ($k = 3$), researcher ($k = 4$), and teaching artist ($k = 3$) under the FE assumptions, $Q(2) = 9.07$, $p < .01$ but do not significantly vary under RE assumptions $Q(2) = 3.96$, $p > .05$. Therefore, I conducted analysis under FE assumptions only. The largest adjusted average effect estimate was for interventions lead by the researcher, (FE: $d = .65$, 95% CI = .29, 1.01). Adjusted average effect estimates for interventions lead by the researcher were significantly different than intervention lead by the classroom teacher under FE assumptions (FE: $d = .01$, 95% CI = -.37, .38, $Q(1) = 5.84$, $p < .01$) as well as significantly different than interventions lead by teaching artists (FE: $d = -.06$, 95% CI = -.42, .29, $Q(1) = 7.59$, $p < .01$). However,

interventions lead by classroom teachers and teaching artists did not vary significantly, $Q(1) = .07$.

Proximity between DBI and measured outcome. Studies were coded for the proximal alignment between the DBI intervention and the outcome that was measured. This sample of studies included 4 that were directly aligned and 6 that were indirectly aligned. All studies were included in the analysis. The average weighted effect of DBI on 21st century skills did not significantly vary under fixed-error assumptions, $Q(1) = .97$, therefore no further analysis of these effects was conducted. The adjusted average weighted effect of DBI on 21st century skills did not significantly vary for proximal alignment between directly aligned ($k = 4$) and indirectly aligned ($k = 6$) under the FE assumptions, $Q(1) = .90$, therefore no additional analyses were conducted.

Domain of DBI intervention. Studies reported effects of DBI on 21st century skills in various domains including: math ($k = 2$), language arts ($k = 2$), and theatre ($k = 2$). The average weighted effect of DBI on 21st century skills significantly varied among the domains of the outcomes under fixed effects, $Q(2) = 23.96$, $p < .001$ as well as under random effects models, $Q(1) = 22.38$, $p < .001$. The largest effect was for language arts (FE: 2.77, 95% CI = 1.61, 3.93; RE: 2.77, 95% CI = 1.61, 3.93). The average weighted effect for language arts significantly differed for all other reported domains under both

fixed- and random-error models including: general academics (FE: .43, 95% CI = .14, .71, $Q(1) = 14.76$, $p < .001$; RE: .47, 95% CI = -.14, 1.08, $Q(1) = 11.86$, $p < .001$), for math (FE: -.03, 95% CI = -.51, .45, $Q(1) = 19.12$, $p < .001$; RE: -.03, 95% CI = -.51, .45, $Q(1) = 19.12$, $p < .001$), as well as theatre (FE: -.45, 95% CI = -1.05, .14, $Q(1) = 23.52$, $p < .001$; RE: -.47, 95% CI = -1.28, .33, $Q(1) = 20.24$, $p < .001$). Interventions in the domains of math and theatre did not significantly differ, $Q(1) = 1.20$.

Studies reported adjusted effects of DBI on 21st century skills in various domains including: math ($k = 2$), and language arts ($k = 3$). The adjusted average weighted effect of DBI on 21st century skills did not significantly vary among the domains of the outcomes under fixed effects, $Q(1) = 2.66$. No further analysis was conducted.

Publication status. All but one of the studies were dissertations; therefore, I could not conduct this moderator analysis.

Measure of outcome. The data presented four categories for measures of the outcomes assessing the effects of DBI on 21st century skills: test ($k = 3$), standardized test ($k = 2$), and survey ($k = 4$). A significant amount of variability was attributed to this moderator under the fixed-error assumptions ($Q(2) = 5.98$, $p < .05$) but not under random-error assumptions ($Q(2) = 3.25$, $p > .05$; therefore, additional analyses were conducted under a fixed-error model only. The largest average weighted effect estimate

of DBI on 21st century skills was measured by a test (FE: d = .84, CI: .18, 1.50, p < .01). This average effect estimate was not significantly different from outcomes measured by standardized tests (FE: d = .17, CI: -.17, .52, p > .05, Q(1) = 3.08, p > .05) but was significantly different from outcomes measured by surveys (FE: d = -.07, CI: -.40, .26, p > .05, Q(1) = 5.90, p < .01). Weighted average effect estimates measured by standardized tests and surveys did not significantly differ from one another (FE: Q(1) = 1.00, p > .05).

Similar results were produced for adjusted weighted effect estimates of DBI on achievement outcomes. Categories included: test (k = 3), standardized test (k = 2), and survey (k = 4) and the moderator significantly accounted for variability under both fixed- and random-error assumptions (FE: Q(3) = 28.45, p < .001; RE: Q(3) = 28.45, p < .001); therefore, additional analyses were conducted with both models. The largest adjusted average weighted effect estimate of DBI on 21st century skills was measured by a test (FE: d = 1.43, CI: .93, 1.94, p < .001; RE: d = 1.43, CI: .93, 1.94, p < .001). This adjusted average effect estimate was significantly different from outcomes measured by standardized tests (FE: d = .24, CI: -.12, .59, p > .05, Q(1) = 14.59, p < .001; RE: d = .24, CI: -.12, .59, p > .05, Q(1) = 14.59, p < .001) and significantly different from outcomes measured by surveys (FE: d = -.21, CI: -.53, .12, p > .05, Q(1) = 28.41, p < .001; RE: d = -.21, CI: -.53, .12, p > .05, Q(1) = 28.41, p < .001). Adjusted weighted average effect

estimates measured by standardized tests and surveys did not significantly differ from one another (FE: $Q(1) = 3.21$, $p > .05$; RE: $Q(1) = 3.21$, $p > .05$).

Type of outcome. In this sample of studies, two categories for type of outcome were present, including: creativity ($k = 5$) and critical thinking ($k = 8$). This moderator did not account for a significant amount of variability in the outcomes under fixed-error assumptions ($Q(1) = .58$, $p > .05$). No further analyses were conducted. For the adjusted average weighted effect estimates, the data presented two categories: creativity ($k = 5$) and critical thinking ($k = 6$). This moderator did not account for a significant amount of variability in the outcomes under fixed-error assumptions ($Q(1) = .84$, $p > .05$). No further analyses were conducted.

Motivation Outcomes Moderator Analysis

Studies that measured motivation included outcomes such as engagement in the process and time on task (for examples, Warner, 2004; Laurin, 2010, respectively). I conducted moderator analyses of the effect of DBI on motivation using five of the moderators of theoretical and methodological interest. Overall, fewer studies reported outcomes measuring motivation than achievement. Therefore, I collapsed categorical divisions when it was theoretically relevant and appropriate. Additionally, only three

studies reported adjusted effects which would cause unstable results; therefore, no analyses were conducted on these effects. Table 11 presents the moderator results.

		Unadjusted *d*-index					
		Fixed			Random		
Moderator	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b
Grade level				9.54**			.24
Lower elementary	3	.17	-.14/.48		.20	-.52/.92	
Upper elementary	2	.39*	.06/.73		.41†	.00/.82	
Gender				5.18			3.60
% female	3	β = -2.20*	-4.10/-.31		β = -2.23†	-4.52/.07	
Number of lesson plans				.35			.17
6 to 10 lessons	2	.15	-.24/.55		.08	-1.04/1.21	
11 to 20 lessons	3	.30*	.01/.59		.33†	-.02/.68	
Leader				1.02			.15
Classroom teacher	2	.39*	.06/.73		.41*	-.00/.82	
Researcher	2	.12	-.33/.57		.20	-1.20/1.61	
Teaching artist	2	.37†	-.02/.76		.51	-.23/1.25	
Proximity between DBI and outcome				1.01			.07
Directly related	2	.12†	-.34/.57		.20†	-1.20/1.61	
Indirectly related	4	.38**	.13/.64		.40**	.11/.70	
Publication status				.27			.00
Dissertation	2	.37†	-.02/.76		.51	-.24/1.25	
Journal	3	.50***	.20/.80		.54**	.14/.94	

Note. CI = confidence interval; DBI = drama-based instruction.

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

The *k* for each moderator may not add up to the overall *k* from Table 6. Studies were excluded for moderator analyses due to insufficient reporting or some studies contributed more than one effect when the study had multiple levels of the moderator.

Table 11. Treatment and student moderators of the effects of DBI on motivational outcomes.

Grade level. First, I examined the association between the magnitude of effect sizes and the grade level of the sample participants. Studies were divided into categories by grade level: lower elementary (k = 3) and upper elementary (k = 2). All studies were included in this analysis. The average weighted effect of DBI on motivation did significantly vary for different grade levels under FE assumptions, $Q(1) = 9.54$, $p < .01$, but not under RE assumptions, $Q(1) = .24$, favoring upper elementary.

Gender. I conducted a regression analysis on the percent of females in a sample for each study. For the 3 samples reporting percent of gender in the sample, the average weighted effect estimate of DBI on achievement was negative and significant under FE assumptions ($\beta = .-2.20$, $p < .05$, $Q(1) = 5.18$) but not significant under RE assumptions ($\beta = .-2.23$, $p > .05$, $Q(1) = 3.60$). No further analysis was conducted. I was unable to conduct analyses on this moderator for adjusted effects due to limited reporting.

Number of lessons. The reported categories for number of lessons were 6-10 lessons (k = 2) and 11-20 lessons (k = 3). One study was excluded from this analysis because the number of lessons was not reported (Enciso & Lee, NP). The average weighted effect of DBI on motivation did not significantly vary due to number of lessons under fixed-error assumptions, $Q(1) = .35$; therefore, no further analysis was conducted.

Leader. Studies reported one of three types of facilitators/leaders for the DBI instruction: the classroom teacher (k = 2), the researcher (k = 2), or the teaching artist (k = 2). One study was excluded from this analysis due to insufficient reporting (Kariuki, 2006). The analysis of the average weighted effect of DBI on motivation revealed no significant variation for different types of leaders under FE assumptions, $Q(2) = 1.02$; therefore, no further analysis was conducted.

Proximity between DBI and measured outcome. Studies were coded for the proximal alignment between the DBI intervention and the outcome that was measured. This sample of studies included 2 that were directly aligned and 4 that were indirectly aligned. All studies were included in the analysis. The average weighted effect of DBI on motivation did not significantly vary under fixed-error assumptions, $Q(1) = 1.01$; therefore, no further analysis was conducted.

Publication status. For this moderator, the reports presented two categories of publication status: dissertations (k = 2) and journals (k = 3). Publication status did not account for a significant amount of the variability in the weighted average estimate effects (FE: $Q(1) = .27$, $p > .05$; RE: $Q(1) = 0$, $p > .05$); therefore, no additional analyses were conducted.

Measure of outcome. Due to insufficient variability, I was unable to conduct a moderator analysis.

Type of outcome. Due to insufficient variability, I was unable to conduct a moderator analysis.

Social Skills Moderator Analysis

Studies that measured social skills included outcomes such as conflict resolution and recognizing emotion in others (for examples, Walsh-Bowers, 1992; Smith, 2010, respectively). I conducted moderator analyses of the effect of DBI on social skills using only two of the moderators of theoretical and methodological interest. Overall, there was not enough variability among the other moderators to do further analysis. Table 12 presents the moderator results.

	Unadjusted *d*-index							Adjusted *d*-index						
		Fixed			Random				Fixed			Random		
Moderator	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b	*k*	*d*	--/+CI	Q_b	*d*	--/+CI	Q_b
Gender				1.10			1.10				.13			.13
% female	3	β = 2.40	-2.08/6.88		β = 2.40	-2.08/6.88		3	β = -.84	-5.31/3.63		β = -.84	-5.31/3.64	
Proximity between DBI and outcome				.03			.03				.00			.00
Directly related	3	.07	-.17/.30		.07	-.18/.32		3	-.02	-.25/.22		-.02	-.25/.22	
Indirectly related	3	.04	-.22/.30		.04	-.22/.30		2	-.01	-.50/.47		-.01	-.50/.47	
Type of outcome				1.35			1.18				.08			.08
With peers	6	.11	-.07/.28		.11	-.07/.28		4	.01	-.24/.25		.01	-.24/.25	
Problem behavior	3	-.07	-.32/.18		-.07	-.35/.20		3	-.04	-.29/.21		-.04	-.29/.21	

Note. CI = confidence interval; DBI = drama-based instruction. †$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.
The *k* for each moderator may not add up to the overall *k* from Table 6. Studies were excluded for moderator analyses due to insufficient reporting or some studies contributed more than one effect when the study had multiple levels of the moderator.

Table 12. Treatment and student moderators of the effects of DBI on social skills outcomes.

Gender. I conducted a regression analysis on the percent of females in a sample for each study. For the 3 samples reporting percent of gender in the sample, the average weighted effect estimate of DBI on achievement was positive, but not significant under FE assumptions ($\beta = 2.40$, $p > .05$, $Q(1) = 1.10$). No further analysis was conducted. For the adjusted average weighted effect of DBI on achievement, percent female was negative, but not significant under FE assumptions ($\beta = -.84$, $p > .05$, $Q(1) = .13$). No further analysis was conducted.

Proximity between DBI and measured outcome. Studies were coded for the proximal alignment between the DBI intervention and the outcome that was measured. This sample of studies included 3 that were directly aligned and 3 that were indirectly aligned. All studies were included in the analysis. The average weighted effect of DBI on social skills did not significantly vary under fixed-error assumptions, $Q(1) = .03$; therefore, no further analysis was conducted. For adjusted the adjusted weighted average of effects, 3 studies were directly aligned while 2 studies were indirectly aligned. There was no significant variation due to this moderator under the fixed-error assumptions, $Q(1) = 0$.

Publication status. All but one of the reports were journal articles; therefore, I was unable to conduct moderator analyses.

Measure of outcome. Due to insufficient variability, I was unable to conduct moderator analyses.

Type of outcome. For this moderator, I was able to look at the comparison of outcomes that measured social skills in peer relationships (k = 6) and social skills as it related to problem behavior (k = 3). This moderator did not account for a singnificant amount of variability in the average weighted effect estimates of socials skills under fixed- or random-error assumptions (FE: $Q(1) = 1.35$, $p > .05$; RE: $Q(1) = 1.18$, $p > .05$); therefore no further analyses were conducted.

For the adjusted weighted average effect estimate, outcomes that measured socials skills in peer relationships (k = 4) and social skills as it related to problem behavior (k = 3) did not vary significantly under a fixed- or random-error model (FE: $Q(1) = .08$, $p > .05$; RE: $Q(1) = .08$, $p > .05$); therefore no further analyses were conducted.

In this chapter, I reviewed the quantitative results for the present meta-analysis on the effects of DBI on student related outcome. In the following chapter, I will discuss these findings, limitations to the research, implications for policy and practice, and future research.

Chapter Five Discussion

In light of the previous research based on the theories of differentiating instruction, self-determination theory, and constructivism, I would expect that DBI has positive impacts of interactive, autonomy-supportive, and co-constructed learning. However, as noted earlier, very few of the studies outlined the intervention strategies and/or aligned them with these theories. Thus, I cannot confidently state that DBI's effects are supported by the theories, but rather use the theories and the findings from this study as the foundation to begin to create a theory of DBI.

In this discussion chapter, I focus on interesting/contradictory/supportive findings in the hopes to develop a more nuanced understanding of the effects of DBI. Due to the design of the reports in this sample, the unadjusted effects should be considered more cautiously than the adjusted effects. That is, the adjusted effects may account for pre-existing differences between the control and intervention groups; therefore, this potentially allows for a less biased effect estimate. When relevant, I highlight when the unadjusted and adjusted effects agree and/ or disagree.

The results of this meta-analysis suggest that DBI has a positive, statistically significant overall effect on achievement, as well as on a number of academic related outcomes including: attitudes toward academics, 21st century skills, motivation, attitudes

towards others, and drama skills. In addition, results for the effect of DBI on self-perception competencies and social skills were in the predicted direction but not statistically significant. However, only the effects of DBI on achievement, attitudes toward academics, and drama skills are positive and significant under both fixed- and random-effects models suggesting more robust and stable findings. When interpreting the adjusted effect size estimates, outcomes in achievement, attitudes toward academics, 21st century skills, and attitudes towards others are significant under a fixed-effect model; while achievement is positive and significant under a random-effects model as well. Undoubtedly, the impact of DBI on achievement cannot be ignored given these robust findings.

In addition to the overall effects, theoretically driven moderator analyses suggest that the strength of relationship between DBI and many of the outcomes is affected by the duration of the intervention, the facilitator of the intervention, the proximity between the intervention and the outcome, and the domain of the outcome. Interventions that are six or more lessons in duration have a stronger impact than those that are shorter. This is most evident in the effect estimates for the achievement outcomes. The facilitator of the intervention also seems to have an impact on the effect of DBI. In particular, the researcher has a stronger impact on achievement and 21st century skills outcomes while

the teaching artist has the strongest impact on attitude and motivational outcomes. For both achievement and attitudes towards academics, DBI has a stronger impact when the intervention is directly aligned to the measured outcome. In addition, DBI has a stronger impact on achievement and 21st century skills in English language arts as compared to other domains.

It is important to note that some of these findings were based on a small number of effect sizes, so it is difficult to place a great deal of confidence in the specific magnitude of the estimated effects. In particular, I caution the interpretation of the moderator effects for motivation and social skills as they are based on very few studies. One study can significantly change the findings when only three or fewer studies are in the set. Further, many of the studies were excluded from specific moderator analyses because of insufficient reporting in the studies. If researchers, policy-makers, and practitioners hope to improve the implementation and effectiveness of DBI, then future research must report specific information vital for dissemination and replication.

FIT OF DATA TO PREDICTIONS: OVERALL

I predicted that there would be a stronger positive effect for students who received DBI as compared to students who did not receive DBI (control conditions) for each of the outcomes. Overall achievement outcomes are positive and significant under fixed- and

random-effects assumptions for both unadjusted and adjusted effect size estimates (d = .43 to d = .55). To translate this finding to a potentially more meaningful metric, the average student who receives DBI as part of the curriculum would score better than 67% to 70% of the students who did not receive DBI as part of the curriculum (see Cooper, 2008 for complete explanation of U3 metric translation of effect size estimates). Based on past meta-analyses and this current finding, DBI should be considered a viable pedagogical method for teachers to raise achievement outcomes alongside other research-based instructional methods. In addition, attitudes towards academics are significant under almost all the assumptions and models except an adjusted effects random-error model—the most conservative estimate (d = .13 to d = .25). Translating to percentages, this suggests that the average student who is a part of the DBI intervention would report more positive attitudes towards academics over 55% to 59% of students who were not part of the DBI intervention. Thus, the meaningful use of DBI in the classroom not only raises student academic achievement, as well as improving students' attitudes towards that academic achievement. In other words, students can learn and enjoy learning when teachers use DBI to teach the curriculum.

Even though all of the overall effects were positive, self-perception competencies and social skills were not significant. The previous meta-analysis that focused on self-

perception competencies also found no significant effect for these types of outcomes (Conrad & Asher, 2001). Previously, there has not been a meta-analysis on social skills; however, DBI has certainly been touted as an effective means to promote pro-social behaviors which this finding suggests may be unfounded (Deasy, 2002). Self-perception competencies and social skills may not affected by DBI interventions and need to be reconsidered as a potential positive outcome. Thus, it seems that this present meta-analysis aligns with previous researchers who suggest that cautious interpretation of past research is imperative to understand the benefits and limits of DBI (Eisner, 1998; Fleming, et al., 2004; Mages, 2008; Wagner, 1998; Winner & Cooper, 2000).

FIT OF DATA TO PREDICTIONS: MODERATORS

Grade level. I predicted that there would be a stronger, positive effect for lower elementary age students as compared to older students. This hypothesis did not hold for any of the outcomes in this sample of studies. In the previous meta-analysis (Kardash & Wright, 1986; Podliozny, 2004), contradictory findings suggested that this moderator may not have a consistent impact on the effects of DBI. Although I predicted that younger students would benefit more as compared to older students, it does not seem that the variability in the effects of DBI can be attributed to grade level. For the achievement and attitudes toward achievement, grade level is inversely related to the average effect

size estimate; however, for 21st century skills, motivation, and self-perception competencies, the grade level of the participants and the outcome are positively related. Of the 45 studies in this meta-analysis, only 15 studies were conducted with middle/high school students and three studies were conducted with college students. Since these studies were spread across the different outcomes, it is necessary to interpret these findings with caution. Research needs to be done with middle/high school and college students in order to clarify grade level as a potential moderator.

Gender. I predicted that DBI would have a stronger, positive effect for female students as compared to male students. Past research suggested contradictory findings for this potential moderator as well. In one previous meta-analysis that mostly looked at academic outcomes, DBI had a significantly more positive effect for females as compared to males; however, in another meta-analysis focused on self-perceptions, DBI had no varying effects for females verses males. In this meta-analysis, as the percent of girls in a sample increased, the effect of DBI on the outcome decreased for all but one of the outcomes. Social skills were the only outcome that DBI had a more positive effect for girls as compared to boys. Again, many of the studies did not report percent gender for the sample. In addition, none of the studies isolated boys and girls for the intervention; thus, many other possible social constructs could be influencing these effects.

Achievement levels. For this exploratory analysis, I did not have a clear prediction. Based on past meta-analyses (Kardash & Wright, Conrad, 1992), I would expect that DBI would have stronger, positive effects for students who are gifted as compared to students who are considered below grade level. Unfortunately, studies did not sufficiently report the achievement level of students and/or most of the studies were done with typically developing students. It is vital for future research to consider achievement level as a potential moderator and to clarify how students of varying abilities may or may not benefit from DBI as an intervention.

Duration of Intervention. I predicted that there would be a stronger positive effect for interventions that included frequent, brief sessions as compared to would have a stronger more positive effect as compared to interventions in which the sessions are infrequent or the intervention as a whole is brief. I coded for duration of the intervention in various ways (i.e., number of lessons, number of weeks, minutes of the intervention); however, I was only able to conduct moderator analyses on number of lessons due to insufficient reporting. Based on past meta-analyses, interventions that last between a few days to a few weeks did not significantly differ but DBI interventions that span twelve weeks to a year or more did have stronger effects (Conrad, 1992; Deasy, 2002). However, Kardash and Wright found that the effect size was inversely related to the minutes per

session, positively related to sessions per week, and no relationship to the span of the DBI treatment (Kardash & Wright, 1986).

In the present meta-analysis, the number of lessons accounted for a significant amount of the variance in achievement outcomes as well as attitude toward achievement. Specifically for achievement, DBI interventions that consisted of six or more lessons (as many as 100 lessons) were significantly more positive than DBI interventions that consisted of five or fewer lessons. Specifically for attitudes toward academics, interventions that consisted of 11-20 lessons were significantly more positive than interventions that consisted of 21 or more lessons. Taking both of these findings together, the data suggests that DBI interventions need to be longer than five lessons, but after twenty lessons, DBI may not have any additional benefits on academic and attitude-related outcomes.

Type of strategy. I predicted that there would be a stronger positive effect for interventions that included more interactive strategies compared to constructive strategies. In addition, there would be a stronger positive effect for interventions that included more constructive strategies as compared to active strategies. Based on Chi's framework for differentiating instruction (2008), I attempted to conduct this exploratory analysis with this sample of studies; however, the reports supplied insufficient

information for coding for the type of strategies used in the intervention. This code, in particular, caused the majority of the disagreements between coders (23.7%). Some reports gave detailed descriptions of the intervention aligned with national standards (Walker, et. al., 2011); however, most of the studies merely mentioned that they used "drama" or "creative drama" with little description of how this was operationalized in the intervention. In addition, I attempted to code for evidence of the Experiential Learning Cycle (Kolb, 1984) and interventions that were aligned with Self-Determination Theory (Ryan & Deci, 2000) to help account for the types of strategies and approaches to DBI that might be in the research. No reports mentioned these theories nor was there specific evidence of their application in this sample of studies.

Leader. I predicted that there would be a stronger positive effect for interventions delivered by a more experienced facilitator. As noted earlier, reports did not consistently document the leader's years of experience in DBI or the leader's exposure to professional development in DBI (the two possible ways that I attempted to capture experience in the coding guide). As a small step toward understanding the potential moderating effects of the facilitator, I did conduct moderator analyses on the type of leader (i.e., researcher, classroom teacher, or teaching artist).

The only past meta-analysis to test this moderator found that the drama specialist (similar to the teaching artist in this study) had a significantly less positive impact on effect estimates (Conrad, 1992). The analyses revealed some interesting similar patterns; however, more research needs to be conducted before considering these patterns or trends conclusive. For achievement and 21st century skills, the researcher and/or classroom teacher accounted for a significant amount of the variability in positive effect size estimates. However, for the attitudes towards academics and motivation outcomes, the teaching artist and/or classroom teacher accounted for a significant amount of the variability in positive effect size estimates. In addition, for self-perception competency skills, the teaching artist had the most positive effect size estimate although not significant. This pattern may suggest that when classroom teachers use DBI to teach curriculum, they have a positive effect on the students, no matter the outcomes. Alternatively, teaching artists may have a positive effect on students when the outcomes address attitudes, self-perceptions, and less so motivation outcomes. This pattern may be suggesting the need for paired instruction between the classroom teacher and teaching artist. Ideally, teachers understand the content and their students, but it may be that the teaching artist piques student interest and motivation toward the curriculum. When coupled with the classroom teacher's expertise in content, the teaching artist may help

impact the attitude and motivation of students. This also suggests that teaching artists may need more extensive training in pedagogical content knowledge and content knowledge to be more effective in the classroom.

Alignment between DBI and outcome. I predicted for this exploratory analysis that there would be stronger positive effects of DBI when the outcome was well-aligned with and more proximal to the intervention. For most of the outcomes captured in this meta-analysis (achievement, attitude toward academics, self-perception competencies, social skills, and 21st century skills), the directly aligned DBI intervention and outcome measure had a more positive impact on the effect size estimate. In other words, if the DBI intervention focused on writing skills and the outcome was a measure of writing skills (directly aligned), then the average effect size estimate was stronger and more positive than if the DBI intervention focused on writing skills and the outcome was a measure of problem solving (indirectly aligned). For achievement and attitudes toward academics, the alignment between DBI and the outcome accounted for a significant amount of variance under the fixed-effects model. If the DBI intervention is focused on writing skills, we would not expect the positive effect on writing skills to immediately effect math achievement outcomes.

Domain of outcome. Thus far, the research literature does not present a pattern upon which to suggest a directional hypothesis for the differing effects of DBI on academic domains. I conducted an exploratory analysis to test academic domain of the outcome as a potential moderator. Across achievement, 21st century skills, and attitudes towards academics, DBI interventions used in English language arts curriculum seem to have the largest impact. The domain of the intervention accounts for a significant amount of variance under fixed- and random-error models for both achievement and 21st century skills. For achievement, language arts and science are the domains that are significantly different from the other domains. However, in 21st century outcomes, only interventions in English language arts are significantly different from other domains. This may be attributable to various possibilities. 1) Much of DBI focuses on stories and role-playing which closely aligns with ELA curriculum. 2) Evident in this research synthesis (ELA outcomes accounted for the majority of effects), DBI is most prevalent in ELA classrooms; thus having time to develop and implement the most effective ways to use DBI for the best outcomes. 3) Typically handbooks for DBI focus on implementation in the ELA classroom (see Heinig, 1992; McCaslin, 1996; Willhelm, 2002). It may be that this finding suggests that the field needs more access to materials that suggest ways to use DBI in other curricular areas.

Publication status. As an additional check on publication bias, I conducted analyses on the publication status of the set of studies for each outcome. The only outcome that presented significant variability among the categories was for self-perception skills; however, the largest effect estimate was for dissertations. If publication bias were present, I would expect to see the largest effects for journals. Therefore, very little, if any, publication bias was present in the set of studies.

Measure of outcome. To test for bias in the type of measurement tool, I conducted a moderator analysis for each outcome. Only the effect estimates of DBI on 21st century outcomes presented significant variability by measurement tool. The difference only existed between tests measuring the outcome and surveys measuring the outcome. This was a unique effect for this outcome and was not present under the random-effects model; therefore, it likely does not present a threat to reliability for this set of studies.

Type of outcome. For this exploratory analysis, I conducted a moderator analysis on the type of outcome for each of the broader categories of outcomes. If significant, then it may suggest that the outcome categories should not have been grouped together. For this set of studies, only the analysis on DBI effects on self-perception skills presented a significant finding. It showed that studies measuring self-discrepancy skills were greater than those measuring all other self-competency skills. However, when looking at the

adjusted effects, this significant difference is not present. Therefore, this suggests that the studies in this category were likely related and appropriately grouped.

LIMITS TO GENERALIZABILITY.

All of the studies included in this research synthesis and meta-analysis were quasi-experimental meaning that each study had two groups that were not randomly assigned to the condition: one that received DBI and one that did not receive DBI. This could translate to biased findings for various reasons: teachers who are already effective may be self-selecting to do the intervention and/or their students may have already been exposed to active learning strategies similar to DBI. However, this type of design is particularly useful when examining interventions intended for applied settings. That being said, we can focus on particular findings. Only results that are significant under fixed- and random-error models can be generalized to the population. Many of my findings were significant only under fixed-error models and therefore should be interpreted with caution.

In this meta-analysis, I performed an exhaustive literature search and conducted trim-and-fill analyses, but undoubtedly some studies published or unpublished were excluded unintentionally. For studies that were included, I did not use an indication of quality research inclusion criteria (see Cooper, 2008 for a discussion of inclusion

criteria). This allows for a more unbiased view of the research literature; however, it may also include reports conducted based on poorly designed research and interventions.

Another limitation is the lack of thorough reporting in the research literature on drama-based instruction. This has been discussed at length by other researchers (Eisner, 1998; Mages, 2008; Wagner, 1998; Winner & Cooper, 2000); however, it is worth mentioning. The field of drama and education needs to make a practice of reporting standard statistical information (i.e., sample sizes, means, and standard deviations for each sample) as well as intervention information (i.e., DBI strategies used, length of the intervention, leader of the intervention). I was unable to use some studies in the analyses due to insufficient reporting.

IMPLICATIONS FOR POLICY AND PRACTICE FOR THE USE OF DBI IN EDUCATIONAL SETTINGS.

As supported by the robust finding for the effects of DBI on achievement and attitudes towards achievement, DBI needs to be considered an effective instructional approach for teachers in the Pre-Kindergarten through college classrooms. These findings have implications for DBI use in educational settings for policy-makers, professional development providers, college professors, and educators. However, it should be noted

that this study is not offering recommendations for traditional arts programming in the schools (e.g., music, visual arts, drama) as that is beyond the scope of this meta-analysis.

Policies that affect a teacher's pedagogy need to include language that supports and encourages drama-based instruction as a pedagogical approach. As the common core standards are developed for nationwide implementation (NGABP, 2010), policy-makers need to include language that supports the use DBI to meet these standards. Professional development providers and college professors of educators need to consider ways to train in-service and pre-service educators to use DBI in the classroom. In particular, it seems it would be particularly beneficial for educators in areas of English language arts and science to have facility with using DBI strategies.

To that end, educators need to use DBI in meaningful ways to positively impact student outcomes. It may be beneficial for educators to partner with teaching artists to help a constellation of student outcomes. Unfortunately at this point, it is difficult to assess the specific strategies that have an impact; however, researchers certainly have a generalized way that they seem to include DBI in the curriculum. Through various strategies, students embody concepts and stories lead by a facilitator. It is my hope that future research will be able to offer more specific ways to use DBI that show positive impact on student outcomes.

FUTURE RESEARCH.

Based on the findings of this meta-analysis, many areas of DBI warrant further research. The effects of DBI on motivation outcomes need to be investigated. Only six studies reported effects on motivation; however, much of the research and advocacy literature in arts integration suggests that motivation is one of the prevailing reasons to use DBI strategies in the classroom. With so few studies, few of the moderators could be tested. However, I did conduct a moderator analysis on grade level and lower elementary and upper elementary effect estimates were significantly different, favoring upper elementary. It might be that as students get older, DBI may have a stronger impact on their motivation. Unfortunately, I did not have enough studies at the middle, high school, or college levels to allow for analysis. With drop-out rates in high schools escalating, DBI might be one way of to redirect and motivate students toward academics.

As stated earlier, the reports in this sample provided insufficient details about the DBI intervention. In future research, this can be approached in various ways including: through better reporting and through studying discrete DBI strategies' effects. Research reports need to include not only basic information about the intervention (e.g., how many lessons, how many weeks, who led the intervention), but also detailed descriptions of sample lessons or strategies that were used. This would allow for researchers to build

upon one another's work and deepen lines of research to try to understand the similarities among interventions that have stronger positive effects on students' outcomes. In addition, this would aid practitioners in understanding specific ways to implement DBI in the classroom.

Further research needs to be conducted on the differing effects of the DBI leader on student outcomes. Initially, a qualitative study documenting the types of interactions and student responses to classroom teachers and teaching artists could be invaluable to understanding how classroom teachers and teaching artists might partner for stronger positive effects for students. Additionally, researchers may want to explore the type of training that might serve training teaching artists and classroom teachers to use DBI in the classroom. To begin this process, researchers need to document the experience level and type of professional development leaders have when they implement DBI in the classroom.

Further research needs to be conducted on the classroom community that may be developed through using DBI. I was surprised to find no studies investigating these types of outcomes. Qualitative research documents the impact of DBI on community factors through extensive ethnographic and case study methods (Enciso, et. al., 2011; Fels, 2009; Neelands, J. 2009; Smith & McKnight, 2009). It may be beneficial to attempt to use this

past qualitative work to develop measures and quantify these potential effects of DBI on the classroom community.

Finally, future research needs to focus on developing a theory of drama-based instruction. Many reports referenced the theory of Multiple Intelligences (Gardner, 1983) or Constructivism (Vygotsky, 1978), but none of the reports in this sample attempted to generate theory about why DBI might be an effective pedagogy for student outcomes. DBI offers a unique educational opportunity for students to embody their learning in a playful and artful environment. For teachers, DBI offers a unique instructional method to engage students in content and co-construct their ideas in the classroom community. Drama-based pedagogy has significant effects on achievement and attitudes toward academics; therefore, the field of drama and education needs to attend to its distinctive contributions to learning rather than solely using theory from other fields. Creating a theory of DBI is beyond the scope of this meta-analysis, but needs to be pursued for the understanding of why DBI has an impact on students in educational settings.

CONCLUSIONS

Drama-based instruction has significant impact on a variety of outcomes. In particular, DBI has significant effect on achievement in various domains, attitudes toward achievement, 21st century skills, and drama skills. These effects are even greater when

certain characteristics are present, for example, when a classroom teacher trained in DBI for achievement outcomes or a teaching artist uses DBI for attitudinal outcomes. However, the effects of DBI should not be over generalized—especially when pursuing funding streams or policy support. As the definitions and measurements stand today, researchers, policy-makers, or practitioners cannot substantiate the claim that DBI will significantly impact a child's self-concept. Researchers, policy-makers, or practitioners cannot substantiate the claim that more (e.g., more lessons, more minutes, more weeks) of DBI equals better effects. As of now, the evidence of this research does not support such statements. These claims discredit the immense measurable impact DBI is having in classrooms across the US. This may not be a catchy slogan like "Art: ask for more" (Americans for the Arts), but this research suggests that there are benefits and limits to DBI. As a researcher and practitioner, I know DBI makes a difference in the academic success of young people. Rather than suggesting that DBI is a panacea for all educational woes, I invite researchers, policy-makers, and practitioners to build upon these findings and pursue research, policy, and practice that facilitate the meaningful, intentional, and effective use of drama-based instruction.

Appendix Coding Guide

Report	
RP1. What was the first author's last name? NOTE: If the same author has multiple reports in one year, place numbers after the name (e.g. Smith 1)	_____ please print neatly ☺
RP2. What was the year of appearance of the report or publication	___ ___ ___ ___
RP3. What was the type of the publication?	___ journal article ___ book chapter ___ book ___ dissertation ___ MA or MFA thesis ___ private report ___ government report (state or federal) ___ school or district report ___ conference paper ___ other; specify: _____

Study Information	
ST1. What was the study number? *(Used to identify reports with multiple studies)* *NO NEED TO FILL THIS IN UNLESS THERE IS MORE THAN ONE REPORT BY THE SAME AUTHOR.*	____
ST2. What type of organization produced this report? Who is driving the research? (Public schools and districts are "government", private schools are "other.")	___ university ___ government entity specify _____ ___ contract research firm specify_____ ___ Private school ___ other specify_____ ___ can't tell
ST3. What was the sampling procedure for recruiting participants in the study? (most will be convenience.)	___ Convenience sample ___ Random selection
ST4. Was this research funded? ST4a. If yes, who was the funder?	___ no ___ yes ___ can't tell ___ federal government specify_____ ___ private foundation specify _____ ___ other specify _____
Coder Name	

Characteristics of the Drama-Based Instruction Intervention	
Fill in as many of the following descriptions as possible.	
I1. How many minutes was the DBI intervention per day?	_____ minutes ___can't tell
I2. How many days was the intervention?	_____days ___can't tell
I3. How often was the intervention?	_____day(s) per month ___can't tell
I4. Overall, how many weeks did the intervention last?	_____weeks ___can't tell
I5. How many lessons were taught?	_____lessons ___can't tell
I6. How many total hours was the intervention? (Compute this if possible.)	_____total hours ___can't tell
Is the treatment facilitated by a leader?	Yes No Can't tell
Are there goals for the treatment? (Alternative would be like free play without any knowledge of where the drama is going.)	Yes No Can't tell

Was the treatment process-oriented? (Alternatively the drama is solely focused on a production.)	Yes No Can't tell
Did the treatment involve any drama strategies?	Yes No Can't tell
I7. What word do they use to describe the intervention?	___Creative drama ___Drama-based instruction ___Arts-Based learning ___applied theatre ___process drama ___arts integration ___enactment strategies ___Image work ___Drama ___Other ___Theatre ___Theatre in education ___drama in education ___Theatre of the oppressed
I8. What theoretical framework did the report used to describe dbi?	___Contructivism/Vygotsky ___Critical Pedagogy/Freire ___Process Drama/Heathcote ___Social Learning/Bruner ___Other; Describe _____ ___N/R
I9. What types of activities do they describe as the drama-based intervention? (Check all that apply.)	___Story dramatization ___Creative drama ___Script writing ___Design-related ___Role play ___Improvisation ___Enactment ___Reflection ___Image work ___Theatre games ___Process drama ___Creative Movement ___Active Discussion

	___Other; Specify_____
Is there a detailed description of the treatment?	None Mentioned activities Lesson plan included Multiple lessons included Full scope and sequence included
I10. Is the intervention linked to state and/or national curricular standards? I10a. If so, standards in which domain(s)? (Choose all that apply.)	___Yes ___No ___Theatre ___Other arts ___Language Arts ___Social Studies ___Math ___Science ___Other arts ___Other; Specify_____
Is the treatment autonomy-supportive for the participants? (Example, students are given choices throughout the drama.)	Yes No Can't tell If yes, give an example:
Is there evidence for the experiential learning cycle? (Example: Leader poses a problem, students reflect on problem, create potential solution, experiment with possible solution, reflect, repeat the process.)	Yes No Can't tell If yes, give an example:
I11. Who led the drama-based instruction? (If more than one, fill in I11-I12.)	___ Classroom teacher ___Theatre teacher ___Researcher ___Teaching Artist/ CD Specialist ___Other arts teacher ___Other

I11a.How much experience does the leader have in drama-based instruction? (Choose all that apply.)	___Can't tell ___ Never done dbi before ___1 workshop/professional development ___Multiple workshops/professional development ___1 college course in dbi ___More than 1 college course in dbi ___Other theatre exposure (major/minor, professional performance) ___Can't tell
I12. Who else led the drama-based instruction? I12a.How much experience does the leader have in drama-based instruction? (Choose all that apply.)	___ Classroom teacher ___Theatre teacher ___Researcher ___Teaching Artist/ CD Specialist ___Other arts teacher ___Can't tell ___ Never done dbi before ___1 workshop/professional development ___Multiple workshops/professional development ___1 college course in dbi ___More than 1 college course in dbi ___Other theatre exposure (major/minor, professional performance) ___Can't tell
I13. Who else led the drama-based instruction? I13a.How much experience does the leader have in drama-based instruction? (Choose all that apply.)	___ Classroom teacher ___Theatre teacher ___Researcher ___Teaching Artist/ CD Specialist ___Other arts teacher ___Can't tell ___ Never done dbi before

	___1 workshop/professional development ___Multiple workshops/professional development ___1 college course in dbi ___More than 1 college course in dbi ___Other theatre exposure (major/minor, professional performance) ___Can't tell
I14. Who else was in the classroom during the drama-based instruction?	___ Classroom teacher ___Theatre teacher ___Researcher ___Teaching Artist/ CD Specialist ___Other arts teacher ___Parents ___N/A
I15. What kind of training did the main leader (from I11) receive **for the dbi intervention**? (Choose all that apply.)	___Written instructions ___Verbal instructions ___Online/video training Specify: _____ ___Face to face training(s) Specify: _____ ___N/R
I16. What were the main leader's characteristics (from I11)? Check all that apply.	____Female ____Male ____White ____Black ____Hispanic ____Asian ____Certified teacher Specify type _____ ____Years teaching
I17. Was there a prescribed scope and sequence for the dbi intervention? (Lesson plans)	___Yes ___No

I18. What domain was drama-based instruction used in? (Check all that apply.)	___ General or mixed academic ___ Reading ___ Language arts ___ Math ___ Social Studies ___ Humanities ___ Natural Science ___ Social Science ___ Sports ___ Social/Emotional ___ Music ___Art ___Theatre ___ Other; specify: _____ ___ NR Describe specific topic: _____
I19. Was there a measure of integrity for the treatment? I19a. How did they indicate/measure integrity? I19b.What level of integrity did they indicate?	___Yes ___No ___ Observer rating/Checklist ___Teacher report ___Professional development report ___Other; specify: _____ ___N/R _____

Characteristics of the Control Condition	
C1. Did the control group have "business as usual"?	___Yes ___No
If you answered "yes", skip to the next section on setting characteristics.	
C2. If you answered "no", what did the control group do?	Specify _____
C3. If not, how many minutes was the control treatment per day?	_____ minutes
C4. How many days was the control?	_____days
C5. How often was the control?	_____day(s) per month
C6. Overall, how many weeks did the control last?	_____weeks
C7. How many lessons were taught?	_____lessons ___can't tell
C8. How many total hours was the control?	_____total hours
C9. Who lead the control? (You can indicate more than one.)	___ Classroom teacher ___Theatre teacher ___Researcher ___Teaching Artist/ CD Specialist

C9a. Who else was involved in the control?	___Other arts teacher ___Can't tell ___ Classroom teacher ___Theatre teacher ___Researcher ___Teaching Artist/ CD Specialist ___Other arts teacher ___Parents ___N/A
C10. What domain was control treatment used in? (Check all that apply.)	___ General or mixed academic ___ Reading ___ Language arts ___ Math ___ Social Studies ___ Humanities ___ Natural Science ___ Social Science ___ Sports ___ Social/Emotional ___ Music ___Art ___Theatre ___ Other; specify: _____ ___ NR Describe specific topic: _____

Setting Characteristics	
S1. What state was the study conducted in? (Use postal codes.) S1. A. If another country, other than U.S., what country?	___ ___ _____
S2. What type of community was the study conducted in?	___ urban ___ small city ___ suburban ___ rural ___ multiple types of communities ___ can't tell but school district given specify _____ ___ can't tell
S3. What type of school/classroom was the study conducted in? (check all that apply if more than one site.)	___public school ___private school non religious ___private school with a religious affiliation Specify religious group _____ ___Charter ___Magnet ___Title One ___School rating status ___Exemplary ___Recognized ___Academically Acceptable ___Failing ___Afterschool ___Pull out classroom (e.g., special education, gifted) ___college/university setting ___Nonschool Specify _____ ___other

	Specify _____ ___can't tell

Research Design	
RD1. What was the research design? (We are assuming that this is not a correlational study.)	___ Random assignment ___ Nonequivalent control group with equating (matching) ___ Nonequivalent control group without equating (matching) ___ Cohort Design (students from previous year before compared to current) ___ Other specify_____ ___ can't tell
THIS IS THE ONLY QUESTION ON ATTRITION RD2. Was there evidence that the groups experienced attrition for different reasons? (Note: Attrition is the loss of participants from groups. This question asks specifically about attrition for different reasons not about whether attrition occurred at all. We will calculate attrition rates based on your answers to the four questions on sample size.)	___ no, there was evidence that little attrition occurred ___no, the report says that groups did not experience attrition for different reasons ___yes, the report says that groups experienced attrition for different reasons ___can't tell, the report says nothing about attrition for different reasons
THIS DOES NOT REFER TO ATTRITION RD3. What characteristics of students or schools were used to equate (match) groups? (Place a 1 in each column that applies, 0 otherwise. Note: Equating refers to procedures used to make groups more comparable. Equating can be accomplished by design (e.g., matching, blocking, stratifying) or statistical controls (propensity score matching, covariance analysis). If no matching, mark N/A.	___Pretest of the outcome measure(s) ___ socio-economic status ___ethnicity ___gender ___prior achievement ___achievement motivation ___other specify_____ ___ N/A
RD4. When was equating (matching) of	___ not applicable (there was no equating)

groups done?	___before dbi intervention ___after dbi intervention was over ___both before and after dbi intervention ___can't tell
RD5. Was there evidence that the equating procedure led to the need for differential statistical regression? (Note. For example, you would answer "yes" to this question if students in the dbi group were chosen from poorly-performing students within a high-achieving school and non dbi students were chosen from high-achieving students within a poorly-performing school.)	___no, differential statistical regression is highly unlikely, given the design ___yes, differential statistical regression is likely, given the design ___N/A, there was no equating done ___ the likelihood of differential statistical regression cannot be assessed from the report
RD6. Was there evidence of the occurrence of a local history event? (Note: A local history event is an event that (a) occurred between the beginning of dbi and the administration of the outcome measure (b) affected either the dbi or non dbi groups (not both) and (c) could have produced the observed outcome in the absence of the intervention.	___ no, the report said nothing to suggest the occurrence of a local history event ___yes, the report said a local history event occurred Describe_____
RD7. Was there evidence that another contaminant or contaminants occurred that could have produced the observed outcome in the absence of dbi? (Note: Exclude from this answer the four types of events you have already coded. These are expectancy, novelty, and disruption effects and local history events. Please be sure to describe in the text box.)	___no, the report said nothing about any other contaminant ___yes, the report said that a contaminant occurred Describe_____

Participant and Sample Characteristics Complete these questions separately for each sample within an dbi intervention for which there is a separate outcome.	
P1. What is this sample's ID number? (Assign a number to each independent sample. We assign number 1 unless there is more than one sample in the study.) P1a. Is there a subsample for this group? (Ex. Outcomes for overall sample and outcomes for girls/boys) If yes, assign a letter for the subsample and complete this section for the subsample. (Eg., Overall sample=1; girls subsample=1a; boys subsample=1b)	___ ___Yes ___No ___
P2. Which of the following labels, if any, could be applied to students in this sample? (Check all that apply.)	___gifted ___typically developing ___"at risk" ___underachieving/below grade level/remedial ___students with disabilities ____learning disability ____emotional disorder ____behavior disorder ____autistic spectrum ____EMR ____physical disability ___ELL/ESL ___other specify_____ ___can't tell
P3. What was the socio-economic status of students in the study? (Check all that apply.)	___low SES ___low-middle SES ___middle SES ___middle-upper SES ___upper SES ___only labeled as "mixed"

P3a. Did the report offer specific numbers for free/reduced lunch? If so, what percentage are reported?	___can't tell ___Yes ___No ___% Students receiving free/reduced lunch ___% Students receiving free lunch ___% students receiving reduced lunch
P4. What was the average age of students in the study? P5. What was the median age of students? P6. What was the youngest age of students? P7. What was the oldest age of students?	___ ___ ___ ___ ___NR
P8. What was the education level of the students in the sample? P8a. If students were preschool to 12th graders, what was the grade level of the sample?	___ Preschool ___ Elementary ___ Lower elementary (K-2) ___ Upper elementary (3-5/6) ___ Middle/junior high school students ___ High school students ___ College students ___ Other; specify: _____ . NR ___ Pre-K ___ K ___ 1 ___ 2 ___ 3 ___ 4 ___ 5

	___ 6 ___ 7 ___ 8 ___ 9 ___ 10 ___ 11 ___ 12 ___NA/NR
P9. What student sexes were represented in the sample? (Calculate percentages.)	___ males, specify percentage:_____ ___ females, specify percentage:_____ ___ NR
P10. What race/ethnicities were represented in the sample? (When possible, calculate percentages.)	___ White, specify percentage:_____ ___ Black, specify percentage:_____ ___ Asian, specify percentage:_____ ___ Hispanic, specify percentage:_____ ___ Native American, specify percentage:_____ ___ Other; specify type and percentage: _____ ___ Not-specified mixed ___NR

Outcome Measure Complete these questions separately for each relevant outcome within each sample. Only complete this section for outcome measures that are administered to both control and treatment groups.	
O1. What is this outcome's ID number? (Assign one number for each discrete outcome per sample. If there are subtests, then assign a letter in addition to the number e.g., 1a, 1b. For separate outcomes assign different numbers, e.g., 1, 2.)	___
O2. What type of outcome measure was this? Note: for example attitudes toward reading, check "attitudes" and "reading" below. (Check all that apply.) In the "specify", give the name of the test/measure if there is a published/reported name of the measure. Ex. TAKS test	___achievement ___attitudes specify _____ ___behavior specify_____ ___motivation Specify _____ ___study habits and skills ___arts ___21st century skills (communication, collaboration, creativity)
O2a. What subject matter did this outcome measure? (Check all that apply.)	___ General or mixed academic ___ Reading ___ Language arts ___ Math ___ Social Studies ___ Humanities ___ Natural Science ___ Social Science ___ Sports ___ Music ___Art ___Theatre ___ Other; specify: _____ ___ NR

	Describe: _____
Proximity between the outcome and the treatment. Did the outcome measured directly relate to the treatment given (e.g., treatment in theatre skills and tested theatre skills) or indirectly (e.g., treatment in reading and tested math skills)?	Directly Inidrectly Can't tell
O3. Who provided the responses to the measure? If observers rated the outcome, then put "observers" as respondent.	___ students ___ parents ___ Classroom teacher ___Theatre teacher ___ other school personnel specify _____ ___Researcher ___PD provider/Teaching Artist ___ Observers Specify _____ ___ other specify _____ ___ can't tell
O4. Who was the referent of the measure?	___ students ___ parents ___ Classroom teacher ___Theatre teacher ___ other school personnel specify _____ ___Researcher ___PD provider/Teaching Artist ___ Observers Specify _____ ___ other specify _____ ___ can't tell
O5. How was the outcome measured?	___ standardized achievement test specify _____

O5a. If there was observation data, was there observer/rater training?	___another test measuring achievement (e.g., teacher-developed, textbook chapters) ___ class grades ___ teacher rating of achievement ___Observer/rating scale ___Survey/Scale ___Interview ___School records ___can't tell ___Yes ___No
O6. Who administered the measure?	___Teacher ___Administrator ___School counselor ___Researcher ___Other Specify_____ ___NR
O8. Was evidence presented regarding whether the validity/reliability of this outcome measure reached an acceptable criterion? (Note: A statement indicating that internal consistency was "acceptable" is sufficient, even if the specific value was not reported. A citation to an external source is sufficient.)	___No ___Yes ___ internal consistency ___test-retest correlation ___other specify _____ ___mentioned that reliability and/or validity had been established
O9. What was the sample size of the **dbi group** that completed this outcome measures? # of schools # of classes	 ___ ___ ___ ___ ___ ___

# of teachers # of students	___ ___
O10. What was the sample size of the **control** group that completed this outcome measures? # of schools # of classes # of teachers # of students	___ ___ ___ ___ ___ ___ ___ ___

Effect Size Information	
E1. Can you get a pretest effect size? Was there a pretest measure given? E1a. Can you get an unadjusted post test effect size? Was there a post test measure given? E1b. Can you get an adjusted post test effect size? Were there gains scores (or other adjusted scores) given?	___No ___ Yes ___No ___ Yes ___No ___ Yes
(This is only answered if there is an adjusted effect size.) E2a. How many control variables are reflected in the adjusted effect size? E2b. If there are control variables, which ones are reflected in this ES?	_____ # of control variables ___ Pre test outcome ___pretest of any of the other outcome(s) ___Prior experience in dbi ___ Prior achievement ___ Prior measure of related variable ___ Sex ___ Age/Grade ___ Ethnicity ___ SES ___ Other; specify: _____ .___NA/NR
E3. When was the outcome measure administered relative to the end of the intervention? Check all that apply	_____ prior to treatment End of treatment Delayed after treatment
E4. For this effect size was the unit of assignment the same as the unit of statistical analysis? (Note: For example, if each student was	___ no ___yes ___can't tell

independently selected to go to the dbi group or not, statistical analyses should be conducted using the student as the unit of analysis. However, if entire classrooms were selected to go to dbi group or not, statistical analyses should be conducted using the classroom as the unit of analysis.) E4a. If you answered "no" to E1 was dependence (including dependence arising from clustering) accounted for in estimates of effect size and their standard errors? (Note: Report would need to adjust for standard errors, use HLM or have fixed effects model.)	___ no ___yes ___can't tell
E5. If equating was done post-hoc through a statistical method (i.e. ANCOVA, multiple regression), what covariate(s) were used to equate groups for this outcome? (Check all that apply.)	___pretest of the outcome measure ___pretest of any of the other outcome(s) ___ socio-economic status ___ethnicity ___gender ___prior achievement ___ achievement motivation ___ other specify - _____ ___not applicable
E6. Could the direction of the effect size be identified for this outcome measure? E6a. If yes, what was the direction? (Note: Be sure to be consistent with the direction. E.g., the measure could be attendance or absences.)	___no ___ yes ___-1 = **non dbi group** performed **better** than dbi group ___0 = there was exactly no difference between the groups (actually is 0) ___+1 = **dbi group** performed **better** than non dbi group
E7. Information for effect size computation (fill in everything available, otherwise leave blank):	

M = Mean SD = standard deviation	
DBI group	
Pre-test M of DBI group on outcome	___ ___ ___ . ___ ___
Pre-test SD of DBI group on outcome	___ ___ ___ . ___ ___
Post-test M of DBI group on outcome	___ ___ ___ . ___ ___
Post-test SD of DBI group on outcome	___ ___ ___ . ___ ___
Adjusted M of DBI group on outcome	___ ___ ___ . ___ ___
Adjusted SD of DBI group on outcome	___ ___ ___ . ___ ___
Sample size of DBI group on outcome	___ ___ ___ ___
Control group	
Pre-test M of control group on outcome	___ ___ ___ . ___ ___
Pre-test SD of control group on outcome	___ ___ ___ . ___ ___
Post-test M of control group on outcome	___ ___ ___ . ___ ___
Post-test SD of control group on outcome	___ ___ ___ . ___ ___
Adjusted M of control group on outcome	___ ___ ___ . ___ ___
Adjusted SD of control group on outcome	___ ___ ___ . ___ ___
Sample size of control group on outcome	___ ___ ___ ___
Summary Statistics	
Total sample size	___ ___ ___ ___
F-statistic	___ ___ ___ . ___ ___

Degrees of freedom for F-test	___ ___ ___ ___
p-value from F-test	0. ___ ___
Independent *t*-statistic	___ ___ ___ . ___ ___
Degrees of freedom for t-test	___ ___ ___ ___
p-value from t-test	___. ___ ___
Dependent *t*-statistic	___ ___ ___ . ___ ___
Degrees of freedom for t-test	___ ___ ___ ___
p-value from t-test	___. ___ ___
MS error from an ANCOVA	___ ___ ___ . ___ ___
Degrees of freedom from an ANCOVA	___ ___ ___ ___
Correlation between pre-test (covariate) and post-test of outcome (dependent measure)	___. ___ ___
Dichotomous Outcomes (Chi-square, frequencies, or proportions)	
Chi-square statistic	___ ___ ___ . ___ ___
p-value from Chi-square	___. ___ ___
Frequency of dbi group successes	___ ___ ___ ___
Frequency of non dbi group successes	___ ___ ___ ___
Proportion of dbi group successes	___ ___ ___ ___
Proportion of non dbi group successes	___ ___ ___ ___

E7a. On what page can this information be found in the report?	___ ___ ___ ___
E8a. Could a pretest effect size be derived for this outcome measure? CALCULATE THE ES. If yes, what was the effect size? E8b. Could a posttest effect size be derived for this outcome measure? CALCULATE THE ES. If yes, what was the effect size? E8c. Could an adjusted posttest effect size be derived for this outcome measure? CALCULATE THE ES. E8c. If yes, what was the effect size?	___ no ___yes ___ ___ . ___ ___ ___ no ___yes ___ ___ . ___ ___ ___ no ___yes ___ ___ . ___ ___
E9. If an effect size could be derived, how could it be done? (Note. Choose only one derivation procedure. They are listed in order of preference.)	___ Standard formula (Note: The standard formula for the d-index is the difference between the all-day and half-day kindergarten group means divided by the pooled standard deviation) Given means, sd, and n's. ___Algebraic equivalent of standard formula (Note: This could be a transformation of a t-test, univariate F-test, correlation, or chi-square.) ___Algebraic equivalent of standard formula with imprecise information (e.g., used $p < .05$ to generate an effect size) ___ Nonstandard formula ___No effect size could be derived
E10. For this outcome, were scores roughly normally distributed within groups?	___no ___yes

E11. For this outcome, were variances roughly equivalent across groups (ratio of variances no greater than 3:1)? (Note: square and then compare variances.)	___no ___yes

References

Abdulwahed, M., & Nagy, Z. K. (2009). Applying Kolb's Experiential Learning Cycle for Laboratory Education. *Journal of Engineering Education, 98*(3), 283-294.

Aldredge, P. A. (2010). *The intersection of religion and homosexuality in the social work classroom: A drama-based approach.*

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*(3), 315-343.

Alivernini, F., & Lucidi, F. (2011). Relationship Between Social Context, Self-Efficacy, Motivation, Academic Achievement, and Intention to Drop Out of High School: A Longitudinal Study. *Journal of Educational Research, 104*(4), 241-252. doi: 10.1080/00220671003728062

Barnett, W. S., Jung, K., Yarosz, D. J., Thomas, J., Hornbeck, A., Stechuk, R., & Burns, S. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly, 23*(3), 299-313. doi: 10.1016/j.ecresq.2008.03.001

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497-529.

Beachboard, M., Beachboard, J., Li, W., & Adkison, S. (2011). Cohorts and Relatedness: Self-Determination Theory as an Explanation of How Learning Communities Affect Educational Outcomes. *Research in Higher Education, 52*(8), 853-874. doi: 10.1007/s11162-011-9221-8

Boal, A. (1974). *Theate of the Oppressed* (C. A. Maria-Odilia & L. McBride, Trans.). New York City: Theatre Communications Group.

Bolton, G., Davis, D., & Lawrence, C. (1987). *Gavin Bolton: Selected writings on drama in education*: Longman Group United Kingdon.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). Comprehensive Meta-Analysis (Version 2.1). Englewood, NJ: BioStat.

Bournot-Trites, M., Belliveau, G., Spiliotopoulos, V., & Seror, J. (2007). The role of drama on cultural sensitivity, motivation and literacy in a second language context. *Journal for learning through the arts, 3*(1), 1-35.

Bowell, P., & Heap, B. S. (2001). *Planning Process Drama*. London: David Fulton Publishers, Ltd.

Bramwell, R. (1990). *The effect of drama education on children's attitudes to the elderly and to ageing.* PhD, The University of British Columbia.

Braund, M. R. (1999). Using Drama To Improve Student Teachers' Understanding in the Physical Sciences.

Burman, D. D., Bitan, T., & Booth, J. R. (2008). Sex differences in neural processing of language among children. *Neuropsychologia, 46*(5), 1349-1362.

Burton, B. (2010). Dramatising the Hidden Hurt: Acting against Covert Bullying by Adolescent Girls. *Research in Drama Education, 15*(2), 255-270.

Cawthon, S., & Dawson, K. (2009). Drama for schools: Impact of a drama-based professional development program on teacher self-efficaacy and authentic instruction. *Youth Theatre Journal, 23*, 144-161.

Chi, M. T. H. (2009). Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science, 1*(1), 73-105. doi: 10.1111/j.1756-8765.2008.01005.x

Cho, Y., Weinstein, C. E., & Wicker, F. (2011). Perceived competence and autonomy as moderators of the effects of achievement goal orientations. *Educational Psychology, 31*(4), 393-411. doi: 10.1080/01443410.2011.560597

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Nillsdale, NJ: Lawrence Erlbaum.

Conrad, F. (1992). *The arts in education and a meta-analysis.* PhD, Purdue University, West Lafayette.

Conrad, F., & Asher, J. W. (2000). Self-concept and self-esteem through drama: A meta-analysis. *Youth Theatre Journal, 14*, 78-84.

Cooper, H. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: Sage.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *Handbook of research synthesis and meta-analysis*. New York: Russell Sage.

Davis, J. H., & Evans, M. J. (1982). *Theatre, Children & Youth*. New Orleans: Anchorage Press.

de Jong, W. (2006). From ‘doing’ to ‘knowing what you are doing’: Kolb's learning theory in teaching documentary practice. *Journal of Media Practice, 7*(2), 151-158. doi: 10.1386/jmpr.7.2.151_3

Deasy, J. R. (2002). Critical links: Learning in the arts and student academic and social development. Washington, DC: Arts Education Partnership.

DeCharms, R. (1968). *Personal Causation*. New York: Academic Press.

Deci, E., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E., & Ryan, R. (2008a). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology, 49*, 14-23.

Deci, E., & Ryan, R. (2008b). Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health. *Canadian Psychology, 49*(3), 182-185. doi: 10.1037/a0012801

Dorion, K. R. (2009). Science through Drama: A multiple case exploration of the characteristics of drama activities used in secondary science lessons. *International Journal of Science Education, 31*(16), 2247-2270. doi: 10.1080/09500690802712699

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of American Statistical Association, 95*, 89-98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based bethod of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 276-284.

Education, P. s. C. o. A. (2011). *Reinvesting in arts education: Winning America's future through creative schools*. Washington, D.C.

Eisner, E. W. (1998). Does experience in the arts boost academic achievement? *Art Education, 51*(1), 7.

Enciso, P., Cushman, C., Edmiston, B., Post, R., & Berring, D. (2011). ""Is that what you really want?’: A case study of intracultural ensemble-building within the paradoxes of 'urbanicity’." *Research in Drama Education: the Journal for Applied Theatre and Performance,* 16 (2), 215-234.

Erdman, H. (1991). Conflicts of Interest: Bringing Drama into the Elementary Foreign Language Classroom. *Youth Theatre Journal, 5*(3), 12-14.

Erikson, E. (1959). *Identity and the Life Cycle*. New York: International Universities Press.

Faye, C., & Sharpe, D. (2008). Academic Motivation in University: The Role of Basic Psychological Needs and Identity Formation. *Canadian Journal of Behavioural Science, 40*(4), 189-199. doi: 10.1037/a0012858

Fels, L. (2009). When royalty steps forth: Role drama as an emobodied collective learning system. *Complicity: An International Journal of Complexity and Education.* 6(2), 124-142.

Fleming, M., Merrell, C., & Tymms, P. (2004). The impact of drama on pupils' language, mathematics, and attitudes in two primary schools. *Research in Drama Education, 9*(2), 177-199.

Francis, M. (2007). The impact of drama on pupils' learning in science. *School Science Review, 89*(327), 91-102.

Freeman, G., Sullivan, K., & Fulton, C. R. (2003). Effects of creative drama on self-concept, social skills, and problem behavior. *The Journal of Educational Research, 96*(3), 131-138.

Freire, P. (2007). *Pedagogy of the Oppressed*. New York: Continuum.

Furrer, C., & Skinner, E. (2003). Sense of Relatedness as a Factor in Children's Academic Engagement and Performance. *Journal of Educational Psychology, 95*(1), 148.

Gardner, Howard (1983), *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books.

Gourgey, A. F., Bosseau, J., & Delgado, J. (1984). *The impact of an improvisational dramatics program on school attitude and achievement*. Paper presented at the American Educational Research Association, New Orleans, LA.

Hanley, M. S., & Gay, G. (2002). Teaching Moral Education and Social Action through Drama. *Talking Points, 14*(1), 22-26.

Harris, L. F., & Rosenberg, H. S. (1983). Creative Drama and Affective Response to Literature. *Children's Theatre Review, 32*(2), 21-25.

Harter, S. (1982). The Perceived Competence Scale for Children. *Child Development, 53*(1), 87-97.

Hausmann, L., Schofield, J., & Woods, R. (2007). Sense of Belonging as a Predictor of Intentions to Persist Among African American and White First-Year College Students. *Research in Higher Education, 48*(7), 803-839. doi: 10.1007/s11162-007-9052-9

Heathcote, D., & Bolton, G. (1995). *Drama for learning: Dorothy Heathcote's mantle of the expert approach to education*. Portsmouth: Heinemann.

Hedges, L. B., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Hendrickson, R. H., & Gallegos, F. S. (1972). *Using Creative Dramatics to Improve the English Language Skills of Mexican-American Students. Final Report*. Sonoma, Rohnert Park: California State College.

Heinig, R. B. (1992). *Improvisation with Favorite Tales: Integrating drama into the reading/writing classroom.* Portsmouth, NH: Heinemann.

Herz, B., & Merz, W. (1998). Experiential learning and the effectiveness of economic. *Simulation & Gaming, 29*(2), 238.

Hicks, D. (Ed.). (1996). *Discourse, Learning, and Schooling*. New York City: Cambridge.

Huntsman, K. H. (1982). Improvisational Dramatic Activities: Key to Self-Actualization? *Children's Theatre Review, 31*(2), 3-9.

Hyungshim, J., Reeve, J., & Deci, E. L. (2010). Engaging Students in Learning Activities: It Is Not Autonomy Support or Structure but Autonomy Support and Structure. *Journal of Educational Psychology, 102*(3), 588-600. doi: 10.1037/a0019682

Ingersoll, R. L., & Kase, J. B. (1970). Effects of creative dramatics on learning and retention of classroom material (pp. 1-69). Washington, D. C. : National Center for Educational Research and Development.

Jackson, T. (Ed.). (1993). *Learning through theatre: New perspectives on theatre in education*. London: Routledge.

Karakelle, S. (2009). Enhancing fluent and flexible thinking through the creative drama process. *Thinking Skills and Creativity, 4*, 124-129.

Kardash, C. A. M., & Wright, L. (1986). Does cretaive drama benefit elementary school students: A meta-analysis. *Youth Theatre Journal, 1*, 11-18.

Kariuki, P., & Humphrey, S. G. (2006). *The effects of drama on the performance of at-risk elementary math students*. Paper presented at the Mid-South Educational Research Association, Birmingham, Alabama.

Kase-Polisini, J., & Spector, B. (1994). Improvised drama: A tool for teaching science. *Youth Theatre Journal, 7*(1), 15-19.

Kastens, K. A., & Liben, L. S. (2007). Eliciting self-explanations improves children's performance on a field-based map skills task. *Cognition and Instruction, 25*, 45-74.

Kaul, G., & Pratt, C. (2010). Undergraduate Research Learning Communities for First-Year and Lower-Division Students. *Peer Review, 12*(2), 20-21.

Kayhan, H. C. (2009). Cretive drama in terms of retaining information. *Procedia Social and Behavioral Sciences, 1*, 737-740.

King, A. (1993). From sage on the stage to guide on the side. *College Teaching, 41*(1), 30-35.

Kolb, D. A. (1984). *Experiential Learning: Experience as the soruce of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Kowal, J., & Fortier, M. S. (1999). Motivational Determinants of Flow: Contributions From Self-Determination Theory. *Journal of Social Psychology, 139*(3), 355-368.

Landy, R. J. (1983). The use of distancing in drama therapy. *The Arts in Psychotherapy, 10*(3), 175-185. doi: 10.1016/0197-4556(83)90006-0

Laurin, S. (2010). *The effect of story drama on children's writing skills.* Concordia University, Quebec, Canada.

Lawton, B., & Brandon, P. R. (2005). Descriptive Statistics on Student achievement data collected for the ARTS FIRST Windward research project (Curriuclum Research & Development Group, Trans.): University of Hawai'i at Manoa.

McCaslin, N. (1996). *Creative Drama in the Classroom and Beyond* (Sixth edition ed.). White Plains: Longman Publishers.

Merrell, R. (2005). *The impact of a drama intervention program on the response of the bystander to bullying situations.*

Miller, H., & Rynders, J. E. (1993). Drama: A medium to enhance social interaction between students with and without mental retardation. *Mental Retardation, 31*(4), 228.

Minnaert, A., Boekaerts, M., & De Brabander, C. (2007). Autonomy, competence, and social relatedness in task interest within project-based education. *Psychological Reports, 101*(2), 574-586. doi: 10.2466/pr0.101.2574-586

Miserandino, M. (1996). Children who do well in school: Individual differences in perceived competence and autonomy in. *Journal of Educational Psychology, 88*(2), 203.

Montgomery, K., & Brown, S. (1997). Simulations: Using experiential learning to add relevancy and meaning to introductory courses. *Innovative Higher Education, 21*(3), 217.

Moore, B. H., & Caldwell, H. (1990). The Art of Planning: Drama as Rehearsal for Writing in the Primary Grades. *Youth Theatre Journal, 4*(3), 13-20.

Nagda, B. A., Gurin, P., & Lopez, G. E. (2003). Transformative Pedagogy for Democracy and Social Justice. *Race, Ethnicity & Education, 6*(2), 165.

National Center for Education Statistics. (2009). *Dropout and completion rates in the United States: 2007*. (2009-064). Washington, D.C.: NCES.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). Common Core State Standards. Washington D.C.: National Governors Association Center for Best Practices, Council of Chief State School Officers.

Neelands, J. (2009). Acting together: Ensembles as a democratic process in art and life. *Research in Drama and Education*. *14*(2), 173-189.

Niedermeyer, F., & Oliver, L. (1973). The development of young children's dramatic and public speaking skills. *The elementary school journal, 73*(2), 95-100.

Park, S., Holloway, S., Arendtsz, A., Bempechat, J., & Li, J. (2012). What Makes Students Engaged in Learning? A Time-Use Study of Within- and Between-Individual Predictors of Emotional Engagement in Low-Performing High Schools. *Journal of Youth & Adolescence, 41*(3), 390-401. doi: 10.1007/s10964-011-9738-3

Partnership for 21st Century Skills. (2009). *Framework for 21st century skills*. Washington, D. C.

Perkins, D. N. (1991). What constructivism demands of the learner. *Educational Technology, 31*(9), 19-21.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International University Press.

Pineau, E. L.(1994). Teaching is performance: Reconceptualizing a problematic metaphor. *American Educational Research Journal,* 31(1), 3-25.

Play. (2001). In B. Strickland (Ed.), *The Gale Encyclopedia of Psychology* (2nd ed. ed., pp. 503-504). Detroit: Gale.

Podlozny, A. (2000). Strengthening verbal skills through the use of classroom drama: A clear link. *Journal of Aesthetic Education, 34*(3/4), 239-275.

Poehner, M. E. (2007). Beyond the Test: L2 Dynamic Assessment and the Transcendence of Mediated Learning. *Modern Language Journal, 91*(3), 323-340. doi: 10.1111/j.1540-4781.2007.00583.x

Policy, C. f. E. (2008). *Instructional time in elementary schools: A closer look at changes to specific subjects*. Washington, D.C.

Prendergast, M. (2008). Teacher as performer: Unpacking a metaphor in performance theory and critical performative pedagogy. *International Journal of Education & the Arts, 9*(2). Retrieved 6/6/2013 from http://www.ijea.org/v9n2/.

Roscoe, R. D., & Chi, M. T. H. (2007a). Tutor learning: The role of explaining and responding to questions. *Instructional Science, 36*, 321-350.

Rosenberg, H. S., & et al. (1983). On Quantifying Dramatic Behavior. *Children's Theatre Review, 32*(2), 3-8.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78.

Schechner, R. (1993). *The future of ritual: Writings on culture and performance*. London and New York: Routledge.

Shacker, D. L., & et al. (1993). Using Drama in the French Immersion Program. *Youth Theatre Journal, 8*(1), 3-10.

Shamir, A. (2005). Mediational Teaching Style and Peer Mediation among Junior High School Students. *Education & Society, 23*(2), 63-81.

Shamir, A., & Lazerovitz, T. (2007). Peer mediation intervention for scaffolding self-regulated learning among children with learning disabilities. *European Journal of Special Needs Education, 22*(3), 255-273. doi: 10.1080/08856250701430786

Sloman, K., & Thompson, R. (2010). An Example of Large-group Drama and Cross-year Peer Assessment for Teaching Science in Higher Education. *International Journal of Science Education, 32*(14), 1877-1893. doi: 10.1080/09500690903229312

Smith, H. (2010). *The effects of a drama-based language itnervention on the development of theory of mind and executive functionning.* PhD, Georgia State University.

Smith, R. (1995). The limits and costs of integration in arts education. *Arts Education Policy Review, 96*(5), 21.

Smith, K. & McKnight, K. S. (2009). Remembering to laugh and explore: Improvisational activities for literacy teaching in urban classrooms.International Journal of Education & the Arts, 10(12). Retrieved [date] from http://www.ijea.org/v10n12/.

Spolin, V. (1986). *Theater games for the classroom: A teacher's handbook*. Ebanston, IL: Northwestern University Press.

Stewig, J. W., & Vail, N. J. (1985). The Relation between Creative Drama and Oral Language Growth. *Clearing House, 58*(6), 261-264.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295-312. doi: 10.1016/0959-4752(94)90003-5

Terwel, J., Oers, B. v., Dijk, I. v., & Eeden, P. v. d. (2009). Are representations to be provided or generated in primary methematics education? Effects on transfer. *Educational research and evaluation, 15*(1), 25-44.

Vansteenkiste, M., Smeets, S., Soenens, B., Lens, W., Matos, L., & Deci, E. (2010). Autonomous and controlled regulation of performance-approach goals: Their relations to perfectionism and educational outcomes. *Motivation & Emotion, 34*(4), 333-353. doi: 10.1007/s11031-010-9188-3

Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychology processes*. Cambridge: Harvard University Press.

Wagner, B. J. (1986). *The effects of role playing on written persuasion: An age and channel comparison of fourth and eighth graders.* PhD, University of Illinois at Chicago, Chicago. (47/11-A, 4008)

Wagner, B. J. (1990). Dramatic improvisation in the classroom. In D. I. Rubin & S. Hynds (Eds.), *Perspective on talk and learning*. Urbana, IL: National Council of Teachers of English.

Walker, E., Tabone, C., & Weltsek, G. (2011). When achievement data meet drama and arts integration. *Language Arts, 88*(5), 365-372.

Walsh-Bowers, R., & Basso, R. (1999). Improving early adolescents' peer relations through classroom creative drama: An integrated approach. *Social Work in Education, 21*(1), 23-32.

Ward, W. (Ed.). (1986). *Stories to Dramatize* (6th ed.). New Orleans: Anchorage Press.

Warner, C. D., & Andersen, C. (2004). "Snails are Science": Creating context for science inquiry and writing through process drama. *Youth Theatre Journal, 18*, 68-83.

Webb, N., & Mastergeorge, A. (2003). Promoting effective helping behavior in peer-directed groups. *International Journal of Educational Research, 39*(1-2), 73-97.

Wells, G. (2007). Semiotic Mediation, Dialogue and the Construction of Knowledge. *Human Development (0018716X), 50*(5), 244-274. doi: 10.1159/000106414

Willhelm, J. (2002). *Action strategies for deepening comprehension: Role plays, text structure tableaux, talking statues, and other enrichment techniques that engage students with text*. New York: Scholastic.

Williams, J. J., & Lombrozo, T. (2010). The Role of Explanation in Discovery and Generalization: Evidence From Category Learning. *Cognitive Science, 34*(5), 776-806. doi: 10.1111/j.1551-6709.2010.01113.x

Winner, E., & Cooper, M. (2000). Mute those claims: No evidence (yet) for a causal link between arts study and academic achievement. *Journal of Aesthetic Education, 34*(3/4), 11-75.

Wright, P. R. (2006). Drama education and development of self: Myth or reality? *Social Pscyhology of Education, 9*, 43-65.

Yager, R., & Akcay, N. (2008). Comparison of student learning outcomes in middle school science classes with as STS approach and a typical textbook dominated approach. *Research in middle level education online, 31*(7), 1-16.

Yeh, Y.-C., & Li, M.-L. (2008). Age, Emotion Regulation Strategies, Temperament, Creative Drama, and Preschoolers' Creativity. *Journal of Creative Behavior, 42*(2), 131-148.

Zanitsch, J. (2009). Playing in the Margins: Process Drama as a Prereading Strategy with LGBT YA Literature. *English Journal*, *98*(4), 85-91.